Mental Capacity Law and the Justification of Actions against a Person's Expressed Wishes

A thesis submitted to The University of Manchester for the degree of

Doctor of Laws

in the Faculty of Humanities.

2018

Paul Skowron

School of Law

Table of contents

Word Count for main text,* including all footnotes: 77,702

**Chapters one to eight.*

Abstract

When should it be permissible to act against someone's expressed wishes in their best interests? In both political philosophy and legal practice, answers to this question often appeal to the concept of autonomy. Broadly, the idea is that if a person is sufficiently self-governing, then their wishes must prevail; but if they are not, then their wishes need not be respected when promoting whatever is good for them.

This thesis analyses both philosophical models of autonomy and the practice of judges in England and Wales when implementing the Mental Capacity Act 2005. With regard to the philosophical models, it finds that, despite claims to the contrary, they do not offer a plausible way of assessing whether someone else is autonomous without appealing to values that are not the person's own. With regard to legal practice, it finds that, although judges speak about 'autonomy' in contradictory ways, a coherent account of when they will find that they must respect a person's expressed wishes can be constructed.

This first stage of analysis makes a gulf between 'autonomy' in philosophy and law obvious. When philosophers talk about 'autonomy', they are largely concerned with the person's relationship to themselves. When judges talk about 'autonomy', they are largely concerned with the person's relationship to the world. 'Autonomy' in the philosophical sense cannot justify current practice because it does not deal with the same subject matter.

Analysis of mental capacity cases does, however, allow the development of an alternative justification for actions against a person's expressed wishes. This justification lies in an evaluation of the entire situation, not of the person. It is not reducible to any model of autonomy, not even 'relational' models. Taken seriously, this justification requires a reorientation of the ethics of mental capacity law: away from overreliance on relatively few abstract 'principles' and towards articulating the difficulty and complexity of real situations. The thesis offers two papers towards the development of this latter mode.

Declaration and copyright statement

No portion of the work in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=24420), in any relevant Thesis restriction declarations deposited in the University Library, The University Library's regulations (see http://www.library.manchester.ac.uk/about/regulations/) and in The University's policy on Presentation of Theses.

Table of cases

Evans v UK (2008) 46 EHRR 34 (GC)

F (Mental Patient: Sterilisation), Re [1991] UKHL 1, [1990] 2 AC 1

GW, Re [2014] EWCOP B23

Hospital NHS Trust, A v CD [2015] EWCOP 74, [2016] COPLR 1

Hospital Trust, The v Miss V [2017] EWCOP 20

IM v LM [2014] EWCA Civ 37, [2015] Fam 61

Independent News and Media v A [2010] EWCA Civ 343, [2010] 1 WLR 2262

ITW v Z [2009] EWHC 2525 (Fam), [2011] 1 WLR 344

K v LBX [2012] EWCA Civ 79, [2012] 1 FCR 441

Kicks v Leigh [2014] EWHC 3926 (Ch), [2015] 4 All ER 329

Kings College Hospital NHS Foundation Trust v C [2015] EWCOP 80, [2016] COPLR 50

Local Authority, A v DL [2011] EWHC 1022 (Fam), (2011) 14 CCL Rep 441

Local Authority, A v M [2014] EWCOP 33, [2015] COPLR 6

Local Authority, A v SA [2005] EWHC 2942 (Fam), [2006] 1 FLR 867

London Borough of Islington v QR [2014] EWCOP 26, (2014) 17 CCL Rep 344

London Borough of Redbridge v G [2014] EWCOP 17

London Borough of Southwark v KA [2016] EWCOP 20, [2016] COPLR 461

London Borough of Tower Hamlets v TB [2014] EWCOP 53, [2015] 2 FCR 264

London Borough, A v VT [2011] EWCOP 3806

MB, Re [1997] EWCA Civ 3093, [1997] 2 FCR 541

Mental Health Trust, The v DD (no 2) [2014] EWCOP 13, 142 BMLR 156

Mental Health Trust, The v DD (no 3) [2014] EWCOP 44

Mental Health Trust, The v DD (no 4) [2015] EWCOP 4

MN, Re [2015] EWCA Civ 411, [2016] Fam 87

Montgomery v Lanarkshire Health Board [2015] UKSC 11, [2015] AC 1430

Ms B v An NHS Hospital Trust [2002] EWHC 429 (Fam), [2002] 2 All ER 449

N (An Adult) (Court of Protection: Jurisdiction), Re [2017] UKSC 22, [2017] AC 549

NCC v PB [2014] EWCOP 14, [2015] COPLR 118

NHS Acute Trust, The v C [2016] EWCOP 17

NHS Foundation Trust, A v Ms X [2014] EWCOP 35, [2015] COPLR 11

NHS Foundation Trust v QZ [2017] EWCOP 11

NHS Trust, An v CS [2016] EWCOP 10, [2016] COPLR 187

NHS Trust, An v DJ [2012] EWCOP 3524

NHS Trust, A v Dr A [2013] EWHC 2442 (COP), [2014] Fam 161

Northamptonshire Healthcare NHS Foundation Trust v ML [2014] EWCOP 2, [2014] COPLR 439

Nottinghamshire Healthcare NHS Trust v RC [2014] EWCOP 1317, [2014] COPLR 468

NRA, Re [2015] EWCOP 59, [2015] COPLR 690

OH v Craven [2016] EWHC 3146 (QB)

P (Statutory Will), Re [2009] EWHC 163 (Ch), [2010] Ch 33

PC v City of York Council [2013] EWCA Civ 478, [2014] 2 WLR 1

PL, Re [2015] EWCOP 44

Pretty v UK (2002) 35 EHRR 1

R v A [2014] EWCA Crim 299, [2014] 1 WLR 2469

R v Cooper [2009] UKHL 42, [2010] Crim LR 75

R (Burke) v GMC [2005] EWCA Civ 1003, [2006] QB 273

R (Nicklinson) v MoJ [2014] UKSC 38, [2014] 3 WLR 200

Royal Bank of Scotland v Etridge (No 2) [2001] UKHL 44, [2002] 2 AC 773

S (Child as parent: Adoption: Consent), Re [2017] EWHC 2729 (Fam)

S (Protected Persons), Re [2010] 1 WLR 1082 (COP)

SAD v SED [2017] EWCOP 3

Sidaway v Board of Governors of the Bethlem Royal Hospital [1985] UKHL 1, [1985] AC 871

SL, Re [2017] EWCOP 5

Somerset County Council v MK [2014] EWCOP B25

T (Adult: Refusal of Treatment), Re [1992] EWCA Civ 18, [1993] Fam 95

Thompson v Foy [2009] EWHC 1076 (Ch), [2010] 1 P&CR 16

V v Associated Newspapers [2016] EWCOP 21, [2016] COPLR 236

W v M [2011] EWCOP 2443, [2012] 1 WLR 1653

Walker, Re [2014] EWHC 71 (Ch), [2015] COPLR 348

WBC (Local Authority) v C [2016] EWCOP 4

Westminster City Council v C [2008] EWCA Civ 198, [2009] Fam 11

Westminster City Council v Sykes [2014] EWHC B9 (COP), (2014) 17 CCL Rep 139

Wye Valley NHS Trust v B [2015] EWCOP 60, [2015] COPLR 843

XCC v AA [2012] EWHC 2183 (COP), [2013] 2 All ER 988

Table of legislation

Domestic legislation

Care Act 2014

Family Law Act 1996

Mental Capacity Act 2005

Mental Health Act 1983

Sexual Offences Act 2003

Statutory instruments

Care and Support (Eligibility Criteria) Regulations 2014, SI 2015/313

Care and Support (Independent Advocacy Support) Regulations 2014, SI 2014/2824

Mental Capacity Act 2005 (Independent Mental Capacity Advocates) (General) Regulations 2006, SI 2006/1832

Parliamentary debate and parliamentary reports

HL Deb 10 January 2005, vol 668, col 12 - 13 (Lord Falconer LC)

Joint Committee on the Draft Mental Incapacity Bill, *First Report* (2002-03, HL 189-II, HC 1083-II)

National Audit Office, *Care Act First-Phase Reforms* (2015-16, HC 82)

Select Committee on the Mental Capacity Act 2005, *Mental Capacity Act 2005: Post-Legislative Scrutiny* (2013-14, HL Paper 139)

International law and communications of international legal bodies

Convention for the Protection of Human Rights and Fundamental Freedoms (European Convention on Human Rights, as amended)

Convention on the Rights of Persons with Disabilities (adopted 13 December 2006, entered into force 3 May 2008) 2515 UNTS 3

UN Committee on the Rights of Persons with Disabilities, 'General Comment No 1' (2014) UN Doc CRPD/C/GC/1

Universal Declaration of Human Rights (adopted 10 December 1948) UNGA Res 217 A(III)

Legislation from other jurisdictions

An Act Respecting Fugitives from Justice, and Persons Escaping from the Service of their Masters 1 Stat 302 1793 (USA)

Mental Capacity Act (Northern Ireland) 2016

In memory of my grandparents:

John McGarry

Meg McGarry

Martin Skowron

Mary Skowron

The author

Paul Skowron first became interested in mental capacity law when training as a nurse on a neurological rehabilitation ward. Once qualified, he worked as a staff nurse on a High Dependency Unit, where he found that capacity issues often raised a complicated mixture of practical, legal, and philosophical problems.

He subsequently obtained an LLB (Hons) from the University of Salford, and was then awarded the President's Doctoral Scholarship Award by The University of Manchester to complete this work.

He is now a research associate on the *Mental Health and Justice* initiative, funded by the Wellcome Trust.

Phd Publications

Skowron P, 'The Implications of Meno's Paradox for the Mental Capacity Act 2005' (2016) 24 Medical Law Review 279.

Skowron P, 'Humility when Responding to the Abuse of Adults with Mental Disabilities' (2017) 53 International Journal of Law and Psychiatry 102.

Skowron P, 'The Relationship between Autonomy and Adult Mental Capacity in the Law of England and Wales' (under review by Medical Law Review).

Other Publications

Skowron P, 'Evidence and Causation in Mental Capacity Assessments' (2014) 22 Medical Law Review 631.

Skowron P, 'John Harrington, Towards a Rhetoric of Medical Law' (Book Review) (2017) Medical Law Review doi.org/10.1093/medlaw/fwx048.

Acknowledgments

Thank you to my supervisors, Kirsty Keywood and Nicola Glover-Thomas, for providing encouragement, criticism, and guidance. Beyond that, thank you for spotting opportunities for me, encouraging me take them, and generally giving me some introduction to how the strange world of academia operates. Thank you, perhaps most of all, for being unfailing sources of humour and common sense. It is probably dangerous to attempt a PhD without access to those. Thank you, too, for reading the strange ramblings I occasionally fired your way as I tried to articulate some half-formed idea. Which reminds me, thank you for your patience.

Without the President's Doctoral Scholarship from the University of Manchester this work could not have happened. Thank you to all involved in its provision. Many of the academic staff from the School of Law have either given me feedback on chapters from this thesis or have taken the time to discuss related issues with me. Thank you especially to Rebecca Bennett, Iain Brassington, Margot Brazier, Neil Allen, Alex Mullock, and Annie Austin. When it comes to practical matters, I, like every postgraduate in the School of Law, have habitually gone to Jackie Boardman with my administrative problems. Thank you for always knowing who to contact and what to do.

Thank you to my 'Principles of Law, Medicine and Ethics' seminar groups, and my occasional Master's groups, for keeping me on my toes. Thanks, especially, to the School of Law's community of PhD students for comradeship. In particular, thank you to Laura Pritchard-Jones and Beverly Clough for long conversations about mental capacity law, too much coffee, and genuine support. Thanks to Jonathan Davies, Elizabeth Cook, and Wiebke Greef: especially for PhD curry nights. For all the argument and chatter, thanks to Rosemary Toll, Max Dyck, Elena Zharikova, David Lawrence, Sacha Waxman, Jenny Krutzinna, Amin Al-Astewani, Louise Bouic, David Gibson, Patrick Sullivan, Thibault Moulin, William Floodgate, Dunja Begovic, Viki Hooton, Elizabeth O'Loughlin, Joe Tomlinson, Naomi-Ellen Speechley, Caroline Henagan, and Ran Lu. If I have missed anyone's name, I am sorry.

Further afield, thank you to the practitioners and academics from other institutions who have encouraged me, provided feedback on early drafts, or just sat and chatted about these issues. Peter Bartlett, Lucy Series, Amanda Keeling, George Szmukler, Judy Laing, Alex Ruck-Keene, and Jonathan Fisk all deserve special mention.

Thanks to Jenny Haselwood and Shabab Rizvi for good food and better company; and to Steve and Janet Dean for Kendal, the climbing, and the certainty that a pint is never more than a week away.

Thank you Mum, Dad, Andy, and Stevie for more than I can list.

Thanks Michelle, for being you.

Chapter One

Introduction

This thesis is a long, complicated answer to a short, simple question: when is it permissible to act against someone's expressed wishes in their best interests? Or, to ask the same question in everyday language, when can we fail to do what someone wants because we are doing what they need? The rest of the thesis uses the language of 'expressed wishes'. It is more precise. There is, however, a reason for putting the question more bluntly here, at the start. A side-effect of specialist languages – whether academic, legal, or in health and social care – can be that the immediacy of a hard moral question is lost in a mist of technical issues and professional politeness. In contrast, asking when we can put aside what someone wants for what they need is uncomfortable. It raises an image the person responding 'who are *you* to say what I need?' That is a very good question too.

The simple question about acting against someone's wishes has a complicated answer, in part, because it is ambiguous. It could mean 'when can you legally act against a person's expressed wishes in their best interests?', but it could also mean 'when should you be able to act against a person's expressed wishes in their best interests?' Both of these questions are sometimes given a deceptively straightforward answer. On the legal side, it is sometimes asserted that a person's wishes must be respected if they have mental capacity and need not be respected if they do not have capacity. There is an element of truth to this, but it is an oversimplification, as chapter three shows. On the ethical side, it is sometimes asserted that a person's wishes must be respected if, and only if, they are autonomous. The thesis shows that this answer is completely inadequate. Any plausible justification for actions against a person's expressed wishes cannot draw primarily on autonomy, which concerns the person's relationship to themselves. It must be grounded in their relationship to the world and, beyond that, the characteristics of the entire situation.

Because so much weight is sometimes attached to autonomy when it comes to decisions against a person's expressed wishes, chapter two begins the thesis by examining the most popular type of philosophical model of the concept. These 'dual models', in outline, hold that a desire is autonomous if the person would endorse it, or at least not be alienated from

it, on critical reflection. The chapter, after giving a brief history of these models and their growing influence on legal scholarship, questions whether they can really determine when it should be permissible to act against someone's expressed wishes without relying on values other than the person's own. It concludes that they cannot.

Chapter three, which is a paper currently under review for publication, provides the foundation of the legal analysis. It analyses case law from England and Wales to determine when judges will consider a person so autonomous that their expressed wish must be respected. In domestic law, mental capacity is important to this determination; but the chapter shows that judges have given three apparently conflicting accounts of the relationship between mental capacity and 'autonomy' in this legal sense (which should not be confused with the philosophical use of the word). These three apparently contradictory judicial accounts are then synthesised into a coherent one, which outlines when a person's expressed wishes must be respected in domestic law.

Doing the initial philosophical and legal analysis in separate but consecutive chapters underlines the differences between the philosophical and legal uses of the word 'autonomy'. Chapter four develops this point. It shows that dual models of autonomy are primarily concerned with the person's relationship to themselves, but mental capacity law is primarily concerned with the person's relationship to the world. 'Autonomy' in the philosophical sense is not completely irrelevant to the law, it is sometimes a factor in best interests determinations, but it is far from central to legal decisions about the acceptability of deciding against a person's expressed wishes. Given the doubts about dual models raised in chapter two, this chapter then asks whether the law's attention to the person's relationship with the world might form a better basis for justification than philosophy's attention to the person's relationship with themselves. In doing so, it abstracts an alternate justification from practice, one in which actions against a person's expressed wishes are justified only if four conditions are all met: prospective harm; false belief; necessity; and proportionality.

Chapter five begins by addressing a gap in the analysis to that point. Even if dual models cannot and do not justify actions against a person's expressed wishes, some alternative model of autonomy might. The chapter considers two classes of alternative models, under the combined heading of 'relational autonomy', and shows that they cannot escape the problems with dual models raised in chapter two. It then suggests that it is unlikely that

any 'model of autonomy' ever can plausibly play a central role in the justification of actions against a person's wishes. Autonomy is a characteristic of persons, but any plausible justification must be grounded on the evaluation of an entire situation. This sort of broad, situational evaluation is quite different to traditional approaches to the ethics of mental capacity law; but the chapter shows that ethical work in this mode was successfully pioneered by four Oxford philosophers of the Second World War years: Iris Murdoch, Mary Midgley, Elizabeth Anscombe, and Philippa Foot.

The first paragraph of this introduction imagined the person the thesis discusses, the person whose expressed wishes are in question, asking 'who are *you* to say what I need?' Chapters six and seven, which are both published papers, each contribute towards an account of how to answer that question well. Both concern the link between knowledge and action. Chapter six concerns what the person whose wishes are in question must fail to know, if overriding those wishes is to be justified. In essence, it is not enough that they fail to recognise a need that they have. They must also fail to recognise that they have failed to recognise that need. If this is so, however, then anyone proposing to act against a person's expressed wishes for their benefit incurs moral duties that domestic law largely neglects. Chapter seven concerns the person who proposes to act on behalf of another in situations of suspected abuse. It argues that that they must actively attend to their own limits; and suggests that attention to this need for humility is a necessary counterpart to legal discussions of a binary between 'autonomy' and 'protection'.

Chapter eight concludes the thesis with two short discussions: one about its original contributions, and one about future research. Future research is needed. If this thesis is a complicated answer to a simple question, then it is also a partial answer. Apart from mental capacity law and the inherent jurisdiction of the High Court, other areas of law bear on when it is permissible to act against a person's expresses wishes for their own good. Chapter three notes that the Mental Health Act 1983 and law pertaining to children, in particular, further complicate the legal picture. Another area requiring future research is the UN Convention on the Rights of Persons with Disabilities.[1] Although the Convention is discussed in almost every chapter,[2] this thesis is not a systematic treatment of either its

[1] Convention on the Rights of Persons with Disabilities (adopted 13 December 2006, entered into force 3 May 2008) 2515 UNTS 3.

[2] Chapter five being the sole exception.

implications for domestic law or the justifiability of the controversial positions taken by the Committee that oversees it.[3] The analysis in these pages might, however, contribute to that larger work.

Finally, two short notes on terminology. First, this is a thesis 'by publication': chapters three, six and seven are all also papers written for legal academic journals.[4] For this reason, when the word 'autonomy' is used in these chapters, then it is in the legal sense, not the philosophical one. Second, in order to minimise the number of overly long sentences, 'actions against a person's expressed wishes in their best interests' is sometimes compressed to 'actions against a person's expressed wishes'. When 'in their best interests' is not expressly stated, it should be assumed. Nothing here is about any other kind of action against a person's expressed wishes.

[3] UN Committee on the Rights of Persons with Disabilities, 'General Comment No 1' (2014) UN Doc CRPD/C/GC/1.

[4] At the time of writing, chapter three is under review by Medical Law Review. Chapter six appeared as Paul Skowron, 'The Implications of Meno's Paradox for the Mental Capacity Act 2005' (2016) 24 Medical Law Review 279. Chapter seven appeared as Paul Skowron, 'Humility when Responding to the Abuse of Adults with Mental Disabilities' (2017) 53 International Journal of Law and Psychiatry 102.

Chapter Two

Can a dual model of autonomy justify actions against a person's expressed wishes?

1. Introduction

Three interconnected ideas are sometimes found in mental capacity scholarship. First, when someone is autonomous, then no action against their expressed wishes in their best interest is legitimate; but when someone is not autonomous, such an action may be legitimate.[1] Second, a person's autonomy can be assessed in a way that is neutral with regard to different values or conceptions of the good. Third, analytical political philosophy provides models of autonomy that can be used for this sort of neutral assessment. Some authors use these three related ideas to explain the current system.[2] Others use them to argue for changes to the current system.[3] Both sets of authors, however, agree that value-neutral assessment of a person's autonomy is possible, and that finding a person is not autonomous will legitimise actions against their express wishes. This chapter casts doubt on this combination of ideas. It shows that the underlying philosophical models cannot justify actions against a person's expressed wishes without either giving up value-neutrality or reaching extreme conclusions that have not been argued for in the literature.[4]

The particular sort of neutrality that is sometimes claimed for mental capacity law is a form of neutrality towards 'different conceptions of the good'.[5] This is the idea, in Rawls's words, that 'everyone is assured an equal liberty to pursue whatever plan of life he pleases as long as it does not violate what justice demands'.[6] This is one of the most contested ideas in post-

[1] Usually, this assumption leaves open the possibility of following their expressed wishes. It just asserts that it is *legitimate* to act against their wishes if that benefits them.

[2] Section 3 discusses the legal scholarship; but, for instance, see John Coggon, 'Mental Capacity Law, Autonomy, and Best Interests: An Argument for Conceptual and Practical Clarity in the Court of Protection' (2016) 24 Medical Law Review 396.

[3] For instance, Jonathan Herring and Jesse Wall, 'Autonomy, Capacity and Vulnerable Adults: Filling the Gaps in the Mental Capacity Act (2015) 35 Legal Studies 698.

[4] Such as the idea that virtually no-one is autonomous enough to have their wishes respected, see subsection 5.1.

[5] John Rawls, *A Theory of Justice: Revised Edition* (first published 1971, Harvard University Press 1999) 80.

[6] *Ibid*. See also Ronald Dworkin, 'Liberalism' in *A Matter of Principle* (Harvard University Press 1985) 191.

war political philosophy,[7] and some doubt it is even truly coherent;[8] but these wider debates are largely irrelevant to the purposes of this chapter. Whether or not the ideal of neutrality is desirable in the broader context, the most popular philosophical models of autonomy do not provide an avenue for achieving it in mental capacity law. To show this, the chapter takes the principle of neutrality in one of its more precise versions: the idea that decisions which affect someone should be justifiable from their point of view.[9] It is this variant of neutrality that is most visible in commentary on mental capacity law; for instance, in claims that this area of law operates 'to vindicate [the person's] ordering of values',[10] or that best interests assessors must respect an 'authentic expression of P's will'.[11] However intuitively appealing some may find such claims, this chapter shows that they lack a plausible basis.

Claims of value-neutrality draw on the most influential philosophical models of personal autonomy: 'dual models',[12] also known as 'split-level' and 'hierarchical' models.[13] Dual models effectively separate the person into two parts: a shallow agent identical with all of the person's desires, inclinations, and motivations; and a deeper, autonomous, agent identical to only some of their desires, inclinations, and motivations.[14] For instance, a desire might attributed to the deep agent only if the person would endorse it on reflection. If they would not endorse the desire on reflection, then it is still 'their' desire in terms of the shallow agent, but it is not autonomous. In everyday language, they want the thing (shallow agent), but they do not *really* want it (deep agent).[15]

Applied to the question of when it is acceptable to act against a person's expressed wishes, dual models are sometimes used to hold that wishes attributable to the deep agent must be respected, but that those attributable only to the shallow agent need not be. This position is

[7] For an overview, see Stephen Mulhall and Adam Swift, *Liberals and Communitarians* (first published 1992, 2nd ed, Blackwell 1996).
[8] George Sher, *Beyond Neutrality: Perfectionism and Politics* (Cambridge University Press 1997) 43.
[9] *Ibid* 41-42.
[10] Coggon (n 2) 400.
[11] Emma Cave and Jacinta Tan, 'Severe and Enduring Anorexia Nervosa in the Court of Protection in England And Wales' (2017) 23 International Journal of Mental Health and Capacity Law 4.
[12] Michael Garnett, 'Agency and Inner Freedom' (2017) 51 Noûs 3.
[13] Marilyn A Friedman, 'Autonomy and the Split-Level Self' (1986) 24 Southern Journal of Philosophy 19; Gerald Dworkin, *The Theory and Practice of Autonomy* (Cambridge University Press 1988) 13.
[14] Garnett, (n 12) 7.
[15] Sometimes the legal literature refers to this distinction as a difference between first-order (shallow agent) and second order (deep agent) desires. This terminology is best avoided. It oversimplifies an influential paper by Frankfurt that distinguishes between 'second order desires' in general and 'second order volitions' in particular, and only treats the latter as 'essential to being a person'. Harry G Frankfurt, 'Freedom of the Will and the Concept of a Person' (1971) 68 the Journal of Philosophy 5, 10.

an example of 'soft paternalism', which allows actions against a person's expressed wishes only when they are not believed to be voluntary.[16] On this use of dual models, the deep agent acts voluntarily, but the shallow agent does not. If, consequently, the line between the two can be found without relying on other norms, then all that is needed to justify actions against a person's expressed wishes in their best interests is a justification for overriding non-voluntary actions. This is an appealing step. It replaces a hard problem, reconciling individual freedom with coercive political authority, with an argument that some cases of coercion are more apparent than real.[17] For the step to be plausible, however, it must be possible to neutrally evaluate how autonomous someone is. Otherwise, the very assessment itself is not justifiable from the assessed person's point of view. Without neutral assessments, instead of avoiding the hard problem of reconciling individual freedom and coercive power, dual model autonomy will have merely obscured it. This chapter shows that this is what occurs, and that dual models cannot provide a basis for value-neutral assessment that can justify actions against a person's expressed wishes without simultaneously justifying extreme and illiberal political change.

To show that dual models cannot provide neutral criteria for autonomy that can justify acting against a person's wishes in their best interest, it is necessary to take them at their strongest. For this reason, after discussing the history of dual models in section two and their appearance in the legal literature in section three, this chapter largely focusses on John Christman's model.[18] Christman has spent several decades developing what he hopes to be a realistic account,[19] which can distinguish between when it is and is not appropriate to treat a person paternalistically.[20] If his theory fails in practical application, then that is a reason to suspect that other, more abstract, theories also do so. This point about other theories is not, however, left as mere suspicion. Sections five and six treat Christman as the main interlocutor, but show that the other possible answers to the questions that they pose

[16] Dworkin (n 13) 124. The contrast is with 'hard paternalism', which allows actions even against a person's voluntary wishes. Another example of an involuntary action would be, on some accounts, one made under the influence of coercion or deception, (n 13) 14. There is little consistency in terminology; and legal writers, in particular, sometimes use 'involuntariness' to mean *only* the effects of coercion and deception. Chapter three, which is also a paper written for legal audiences, follows this narrower usage.

[17] On the hard problem more generally see Michael Huemer, *The Problem of Political Authority: An Examination of the Right to Coerce and the Duty to Obey* (Palgrave Macmillan 2013) Part I.

[18] John Christman, *The Politics of Persons: Individual Autonomy and Socio-Historical Selves* (Cambridge University Press 2009).

[19] *Ibid* 3-5.

[20] *Ibid* 162.

must also abandon neutrality if they are to justify actions against a person's expressed wishes.

Climacus, one of Kierkegaard's pseudonyms, writes that 'the crux is exactly the relationship of the subject'.[21] In other words, writing objective accounts of personhood may be difficult, but it is never as difficult as usefully relating those accounts to a real human being. This is the central point of this chapter. Although there are abstract problems with dual models,[22] they pale in comparison to the practical issues. In practice, even simple questions have complex ramifications. The questions posed here are very simple. All dual models distinguish between a person's desires in general and those desires that they would identify with, endorse, or not be alienated from on 'reflection'. Section five merely asks something anyone trying to apply such a theory must necessarily ask: what counts as 'reflection'? The section canvasses four possible answers, only one of which can justify actions against a person's expressed wishes without relying on values other than the person's own or reaching the conclusion that virtually nobody is ever autonomous. Then, section six asks another simple question: how should someone's reflection be measured against this standard? Must the person actually reflect, must they merely be able to reflect, or is it enough to decide what the outcome would be if, hypothetically, they did reflect? When combined with the conclusion of section five, section six shows that there is no option that can justify actions against a person's expressed wishes by drawing only on the person's own values and the idea that the deep agent must be respected but the shallow agent need not be.[23] That being so, dual models do not provide a neutral basis for assessing a person's autonomy.

---

[21] Søren Kierkegaard, *Concluding Unscientific Postscript to the Philosophical Crumbs* (first published 1846, Alastair Hannay tr, Cambridge University Press 2009) 33.

[22] Garnett (n 12).

[23] Garnett notes that this second premise, only the deep agent need be respected, is often left implicit. Possibly, it is taken to be uncontroversial. If so, this seems unduly complacent. A full treatment is beyond the scope of this thesis, but an 'alternative' dual model could justify either celebrating alienation or avoiding reflection. Killmister discusses the former possibility. The latter possibility is similar to the Daoist ideal of *wúwéi*, action performed with such absorption that self-consciousness is dissolved. Michael Garnett, 'Taking the Self out of Self-Rule' (2013) 16 Ethical Theory and Moral Practice 21; Suzy Killmister, 'The Woody Allen Puzzle: How "Authentic Alienation" Complicates Autonomy' (2015) 49 Noûs 729; Chad Hansen, 'Wuwei: Taking No Action' in Antonio S Cua ed, *Encyclopaedia of Chinese Philosophy* (Routledge 2003) 786.

2. Dual models of autonomy

A detailed genealogy of dual models of autonomy would require book-length treatment. Fortunately, it is not necessary here. This thesis is only concerned with their applicability to law, when any such model must address the questions of what standard of reflection is required and how it is to be assessed. Sections five and six show there are simply no plausible answers to these questions that do not rely on other normative concepts, so it is enough if this section traces the common features of dual models and provides historical context for the discussion that follows.

Dual models owe much to the work of John Stuart Mill,[24] and this can seem fortunate. Mill, uniquely among philosophers, is sometimes quoted in the Court of Protection,[25] always for unequivocal affirmations of the right to liberty. Rather less legal attention, however, is paid to why Mill affirmed liberty, to what sort of creatures Mill thought humans must be for liberty to be so important. In other words, Mill's moral psychology is generally neglected by mental capacity lawyers. This is unfortunate, for it is a prototypical dual model. He believes individuals should act from their 'own character' rather than from 'traditions or customs'.[26] Even this opposition implies a contrast between a deep agent who acts for their own reasons and a shallow agent who acts for merely customary reasons, but Mill's dual model is not just implicit. He approvingly quotes Humbolt, 'the end of man …not suggested by vague and transient desires, is the highest and most harmonious development

[24] Sometimes, dual models are referred to as 'Kantian', and this can seem natural. Kant wrote about 'autonomy'. The connection between his work and dual models is, however, weak. As O'Neill points out, autonomy for Kant is willing in accordance with universal law: not willing in accordance with what you desire, would desire on reflection, or should desire. For Kant, treating desire as primary, even the desire to follow the moral law, subordinates the law to the person's own aims, and this is 'radical evil'. In contrast, in dual models, reflection makes desires coherent enough to have binding moral force. For Kant, desire is irrelevant to whether a moral duty exists, but for dual theorists the right kind of desire is everything. Onora O'Neill, 'The Inaugural Address: Autonomy: The Emperor's New Clothes' (2003) 77 Proceedings of the Aristotelian Society, Supplementary Volumes 1; Immanuel Kant, *Groundwork of the Metaphysics of Morals* (first published 1785, Mary Gregor tr, Cambridge University Press 1997) 4:441, 4:459; Immanuel Kant, *Religion Within the Boundaries of Mere Reason* in *Religion and Rational Theology* (first published 1793, Allen W Wood and George Di Giovanni tr, Cambridge University Press 1996) 6:36.

[25] The modern Court of Protection was created by the Mental Capacity Act 2005. Its case law is examined in the next chapter. Cases citing Mill include *Nottinghamshire Healthcare NHS Trust v RC* [2014] EWCOP 1317, [2014] COPLR 468 [8]; *Kings College Hospital NHS Foundation Trust v C* [2015] EWCOP 80, [2016] COPLR 50 [2].

[26] John Stuart Mill, *On Liberty* (first published 1859, David Bromwich and George Kateb eds, Yale University Press 2003) 122.

of his powers to a complete and consistent whole'.[27] Here, the shallow self of 'transient desires' is contrasted with the deep, 'complete and consistent', self in an explicit dual model; and, as with modern dual models, the line between deeper and shallower selves is drawn by personal reflection. Mill distinguishes between those who 'ape-like' let the world choose their 'plan of life' and those who use 'reasoning', 'judgment', and 'discrimination' to decide.[28] This is why Mill thinks political liberty is important: because it will create deeper agents, people who choose their own 'plan of life'.[29]

The 'life plan' strand of philosophy did not die with Mill. It reappears in Royce. For him, the 'central moral maxim' of 'be loyal to loyalty' entails that each person 'inevitably defines for himself a cause'.[30] This, for Royce, means that a self is a 'human life lived according to a plan';[31] and this, too, connects directly to a dual model. If someone lives in this way, they are a 'rational and unified self'; but if they do not, they are merely 'a cauldron of seething and bubbling efforts to be somebody'.[32] Royce's work influenced the most prominent text of post-war political philosophy, Rawls's *Theory of Justice*.[33] Rawls adopts the definition of a person as a human life lived according to plan,[34] and states that 'a rational plan of life establishes the basic point of view from which *all* judgments of value relating to a particular person are to be made'.[35] This is another dual model. It splits the empirical human being into a deep self, living by rational plan, and a shallow shelf, consisting of both the deep self and everything else. For all three of these authors – Mill, Royce, and Rawls – life plan autonomy is a relatively briefly argued premise to a larger political philosophy. In contrast, and more recently, Bratman develops the idea into a detailed model of human agency.[36] His view is more subtle, so harder to briefly summarise. In essence, however, his later position is that a desire can be ascribed to the agent if they would reflectively endorse that desire based on a higher level 'self-governing policy' of treating such desires as providing them

[27] *Ibid* 122, citing Wilhelm von Humboldt, *The Limits of State Action* (first published 1850, JW Burrow tr, Cambridge University Press 1969) 11.
[28] (n 26) 124.
[29] *Ibid* 83.
[30] Josiah Royce, *The Philosophy of Loyalty* (Macmillan 1911) 151.
[31] *Ibid* 168.
[32] *Ibid* 172.
[33] Rawls (n 5).
[34] *Ibid* 358.
[35] *Ibid* 359 (emphasis added). 'Rational' is what the person would choose if they decided 'with full awareness of the relevant facts and after a careful consideration of the consequences', *ibid*.
[36] Michael E Bratman, *Intention, Plans, and Practical Reason* (Harvard University Press 1987).

with reasons for action.[37] This is another dual model. Desires that would not be reflectively endorsed are not considered part of the deep agent.[38]

Bratman is deeply influenced by a second strand of philosophy, one which identifies personhood with the will. It can be argued that similar thoughts occur as far back as Augustine,[39] but the root of the modern philosophical attention is a 1971 article by Frankfurt.[40] For Frankfurt, to will is to have a second-order volition, a desire that another attitude be your effective first-order desire.[41] For example, if it is my will to dance, then I want to desire to dance; *and* I want my desire to dance to prevail over other factors, such as my desire to avoid embarrassing myself. In other words, I want my desire to dance to guide my actions. This is clearly a dual model. Indeed, Frankfurt splits the agent so completely that he is comfortable talking about someone being 'violated by his own desires'.[42] Furthermore, only those that can will in this particular way are, for Frankfurt, persons.[43] This is a position with vast political implications that other writers have directly addressed. For instance, Dworkin, who identifies autonomy with the capacity of an agent to reflect on their desires,[44] uses it as the basis of an extended account of consent in politics and medicine.[45] Similarly, Friedman, who identifies autonomy as a person reflectively 'taking a stance' towards some of their 'wants and values',[46] uses it as the basis of an extended discussion of intimate relationships, domestic violence, and cultural minorities.[47]

This chapter does not challenge all political uses of dual theories; and, more broadly, the dual model literature has been remarkably fertile.[48] The very nature of this work does, however, raise questions about whether dual models have any easy application in the courtroom, ward, or community: 'life plan' autonomy was part of ambitious accounts of, and programs for, the entire liberal order; the work following Frankfurt hoped to describe

[37] Michael E Bratman, *Structures of Agency: Essays* (Oxford University Press 2007) 40.
[38] *Ibid* 4.
[39] Hannah Arendt, *The Life of the Mind* (Harcourt 1971) vol 2 ch 10.
[40] Frankfurt (n 15).
[41] *Ibid* 10-11.
[42] *Ibid* 12.
[43] *Ibid*.
[44] Dworkin (n 13) 15-20.
[45] *Ibid* part II.
[46] Marilyn Friedman, *Autonomy, Gender, Politics* (Oxford University Press 2003) 5.
[47] *Ibid* parts III, IV.
[48] For instance, Korsgaard, who considers the affinity between her work and Frankfurt's 'obvious', develops an account of moral obligation based around reflection on 'practical identity'. Christine M Korsgaard, *The Sources of Normativity* (Onora O'Neill ed, Cambridge University Press 1996) 99.

the fundamental nature of personhood; and even those with a more practical bent, like Dworkin, typically address the relationship between citizen and state on the scale of an entire population. It would be surprising if an idea with this lineage had straightforward practical application to individual cases. This chapter argues that it does not. It does, however, have a connection to legal scholarship.

3. Dual models in legal theory

Two intertwined domains of legal scholarship draw on dual models of autonomy: a medico-legal school primarily concerned with domestic law; and an international human rights school primarily concerned with responding to the UN Convention on the Rights of Persons with Disabilities ('the UNCRPD').[49] To substantiate two claims made in the introduction, this section gives an overview of both literatures. The first claim is that it is often assumed that there is, or should be, a connection between capacity law and dual models of autonomy. The second claim is that it is often assumed that this connection allows the autonomy of an individual to be assessed in a way that is neutral between different conceptions of the good. Legal scholarship's reliance on dual theories, however, has not been rigorous. The two practical questions asked in sections five and six – what counts as reflection and how is to be assessed? – have not been addressed in detail by legal scholars endorsing dual models. In this regard, philosophers have been more conscientious.

The domestic medico-legal school has been influenced by Gerald Dworkin. He, as mentioned in the previous section, uses a dual model in his account of consent in medicine and politics. He does not, however, deal with the practical question of how to assess a particular individual in any depth. Indeed, he assigns it only half a page of his book, saying 'on the level of theory, the situation does not seem too difficult'.[50] Here, he misses the Kierkegaardian point made in the introduction to this chapter: political theory that only stays on its own 'level' never has to deal with the intractable stuff of real human lives, so it never is where the true difficulties are found. Dworkin's book has, nevertheless, influenced domestic medical law scholarship, chiefly through the work of John Coggon. Coggon, in an

[49] Convention on the Rights of Persons with Disabilities (adopted 13 December 2006, entered into force 3 May 2008) 2515 UNTS 3.
[50] Dworkin (n 13) 117. Dworkin's position is a 'mere competence' one. See subsection 6.2.

article that has been influential in turn, distinguishes between three 'kinds' of autonomy.[51] One, 'current desire', is reducible here to a person's expressed wishes; and another, ideal desire, is any concept of autonomy that relies on values other than the person's own.[52] The third, 'best desire autonomy' refers to a dual model.[53] Coggon draws an express connection to the work of Frankfurt and Dworkin, and says that 'on this account, we find a person to be autonomous if he acts in accordance with his own value system'. This is clearly a dual model: actions that 'accord' with the person's value system are attributable to the deep agent, and other actions only to the shallow agent. Furthermore, although Coggon accepts that all three accounts 'have a place in our lives',[54] he believes that the dual model is 'generally the conception to be preferred, at least with regard to serious decision-making'.[55] This is a fairly modest endorsement of dual models, but his position has since hardened. In a more recent article, he refers to best desire autonomy as the 'gold standard' that the courts should be upholding.[56] Other writers have joined Coggon in taking this more extreme position.[57]

For Coggon, mental capacity law is already 'constructed' around this dual model,[58] and this is one form of an assumption discussed in the introduction to this chapter: that there is a connection between 'autonomy' in philosophy and law. Not all of those influenced by Coggon would agree. Herring and Wall endorse his account, but believe that best desire autonomy would require either a 'complete reworking' of the domestic Mental Capacity Act 2005 or a strengthening of the inherent jurisdiction of the High Court.[59] This difference between their position and Coggon's is, however, relatively minor. Although, they do not believe that there is currently a close connection between law and dual models of autonomy, Herring and Wall argue that there can and should be such a connection. Like

[51] John Coggon, 'Varied and Principled Understandings of Autonomy in English Law: Justifiable Inconsistency or Blinkered Moralism?' (2007) 15 Health Care Analysis 235.
[52] *Ibid* 240.
[53] *Ibid* 241. Coggon's distinction between 'ideal' and 'best' desire autonomy seems identical to the philosophical distinction between substantive and procedural accounts of autonomy. For an overview applicable to the current context see: Gareth S Owen and others, 'Mental Capacity and Decisional Autonomy: An Interdisciplinary Challenge (2009) 52 Inquiry 79. See chapter five, subsection 2.1 for further discussion of substantive dual models of autonomy.
[54] (n 51) 244.
[55] *Ibid* 253.
[56] (n 2) 401.
[57] Cave and Tan (n 11).
[58] (n 2) 400.
[59] They favour the latter option. (n 3) 717.

Coggon, they assume that dual models make it possible to assess a person's autonomy while remaining neutral between particular conceptions of the good. On this view, in Coggon's words, the court's task is simply to determine and then act 'in accordance with the values that [the person], on reflection, would endorse'.[60]

A second strand of legal scholarship influenced by dual theories responds to Article 12 of the UNCRPD. This Article requires states to 'recognize that persons with disabilities enjoy legal capacity on an equal basis with others in all aspects of life'; and, importantly, that any 'safeguards' on that right 'respect the rights, will and preferences of the person'.[61] Article 12 was controversial during negotiations, the lines being drawn between those who thought that it should never permit 'substituted decision-making', in which someone else makes a decision in the best interests of a person, and those who thought that such decisions must be allowed as a last resort.[62] The resulting Article does not address the controversy; so, unsurprisingly, it has continued. Indeed, it has escalated. The UN Committee that oversees the Convention has called all existing mental capacity law into question by stating that 'perceived or actual deficits in mental capacity must not be used as justification for denying legal capacity'.[63] Instead, decisions must be made according to the 'best interpretation' of the person's will and preferences.[64] On a superficial reading, this seems to simply assert that a person's expressed wishes must always be respected, but it is not necessarily so straightforward. Some who advocate for this standard's adoption hold that 'a verbal expression in one instance may not necessarily represent the *true* will and preferences of an individual'.[65] In this, there is the outline of a dual theory. The 'verbal expression' is only attributed to the shallow agent, not to the deeper 'true' self. Other responses to the UNCRPD explicitly make this connection. For instance, both Craigie and Szmukler draw on Bratman's dual theory, discussed in the previous section.[66] For Szmukler, this allows a

[60] Coggon (n 2) 401.
[61] UNCRPD Art 12(2),(4).
[62] Amita Dhanda, 'Legal Capacity in the Disability Rights Convention: Stranglehold of the Past or Lodestar for the Future?' (2007) 34 Syracuse Journal of International Law and Commerce 429.
[63] UN Committee on the Rights of Persons with Disabilities, 'General Comment No 1' (2014) UN Doc CRPD/C/GC/1 para 12.
[64] *Ibid* para 21.
[65] Eilionóir Flynn and Anna Arstein-Kerslake, 'Legislating Personhood: Realising the Right to Support in Exercising Legal Capacity' (2014) 10 International Journal of Law in Context 81, 98 (emphasis added). Flynn and Arstein-Kerslake advised the group that drafted the General Comment.
[66] Jillian Craigie, 'Capacity, Value Neutrality and the Ability to Consider the Future' (2013) 9 International Journal of Law in Context 4; George Szmukler, 'The UN Convention on the Rights of Persons with Disabilities:

distinction to be drawn between 'will' and 'preferences'. The former are 'authentic' expressions of 'deep commitment' that demand more respect.[67] Once again, a connection is made between autonomy in political philosophy and in law. Furthermore, when giving examples, this literature often treats the 'authenticity' of a person's preference as a neutral question that can be determined by any interested third party.[68]

The idea that dual models can provide a theoretical basis for value-neutral mental capacity law has been doubted by some authors. Craigie doubts that coercive treatment can be justified by appealing to the idea that a person, when no longer showing the symptoms of mental illness, will reflectively endorse it. When such appeals remain value-neutral, they fail to convincingly show why the person's future self should be favoured over their current self.[69] Similarly, Freyenhagen and O'Shea show that the concept of 'mental disorder' is often used to hide substantive value judgements in what are supposedly value-neutral models of autonomy.[70] More fundamentally, Banner points out that a person's values do not form a fixed point that can be used as the basis of a neutral evaluation: they are often indeterminate and change in response to external factors.[71] Nothing in the rest of this chapter contradicts these critiques, but it uses a different methodology, grounded in the question of application. For any dual model to be practically useful, it must set both a standard and method of evaluating a person's reflection. The process of setting these is fatal to value-neutrality. To show this, however, a dual theory that is both practical-minded and detailed must be examined. John Christman has provided such a model.

4. Christman's model of autonomy

Christman's model is perfect for testing against practice. He believes it is realistic, taking into account the social, historical, and embodied aspects of real humans.[72] He also believes

'Rights, Will and Preferences' in Relation to Mental Health Disabilities' (2017) International Journal of Law and Psychiatry doi.org/10.1016/j.ijlp.2017.06.003.

[67] Szmukler (n 66) 5, although on his account even the deep agent does not demand *complete* respect, 6.

[68] For example: Flynn and Arstein-Kerslake (n 65) 98; Michael Bach and Lana Kerzner, *A New Paradigm for Protecting Autonomy and the Right to Legal Capacity* (Law Commission of Ontario 2010) 144. In contrast, Szmukler denies that any entirely neutral interpretation of another person is ever possible (n 66) 5.

[69] Craigie (n 66).

[70] Fabian Freyenhagen and Tom O'Shea, 'Hidden Substance: Mental Disorder as a Challenge to Normatively Neutral Accounts of Autonomy' (2013) 9 International Journal of Law in Context 53.

[71] Natalie F Banner, 'Can Procedural and Substantive Elements of Decision-making be Reconciled in Assessments of Mental Capacity?' (2013) 9 International Journal of Law in Context 71.

[72] Christman (n 18) 11.

it has 'practical purposes',[73] which include picking out '*all and only* those agents …towards whom paternalistic intervention would be disrespectful'.[74] This is a succinct version of the three ideas discussed in the introduction to this chapter: that a concept of autonomy can draw a distinct line between when it is and is not permissible to act against someone's wishes; that it can do this work alone, without relying on particular conceptions of the good; and that a dual model, his own, provides such a concept. In other words, Christman makes a book-length defence of the position that this chapter contests. This is not the only reason for choosing his work. Christman has an obvious commitment to respecting the variety and fluidity of human life, and has repeatedly modified his model to take criticisms into account. This means that his dual theory avoids some difficulties that others do not. Furthermore, he is relatively precise and unequivocal about what autonomy requires,[75] so his theory can be tested to an extent that vague formulations about people acting 'in accordance with their own values' cannot. In other words, the task of arguing that dual models cannot operate as Christman wishes is made easier by the excellence of his work.

In Christman's account, someone is autonomous with regard to one of their own desires or values if two requirements are met, competence and authenticity.[76] Each of these is specified in some detail. Someone is competent with regard to a motivation if they can form an effective intention on the basis of it and can critically reflect on both it and other 'basic motivating elements' of their 'make-up'.[77] For example, I am competent with regard to my desire to become a nun only if I can form an effective intention to become a nun and I can critically reflect both on that desire and my other relevant motivations, such as my counter-desire to marry. 'Effective' intentions are simply intentions the person can act on. The subsequent action does not need to be successful.[78] If I intend to become a nun, and ask the

[73] *Ibid* 110.
[74] *Ibid* 162 (emphasis added). He does admit that sometimes 'it will not be possible to determine whether conditions defining autonomy are met for particular choices', but never explores the practical implications of this admission, 135.
[75] The following sections do draw attention to some equivocation, but this is relatively minor.
[76] (n 18) 155. At one point Christman indicates a further requirement, 'socially effective open options' (122). This point does not, however, appear in his summary; and when summarising he indicates that constraining circumstances can make someone less 'free' without making them less 'autonomous' (154-155). Arguably, this element is already contained in his requirement that there be no 'constraint' on reflection, discussed below and in subsection 5.3.
[77] *Ibid* 147, 155.
[78] *Ibid* 154.

Mother Superior for permission, but she refuses because I am male, then my intention was nevertheless effective. I acted on it, even if the action was unsuccessful.

Christman's second condition for someone to be autonomous with regard to a motivating characteristic is authenticity. This is specified as someone not judging themselves to be alienated from the characteristic if they were to critically reflect on it in a sustained manner and in light of how it was formed.[79] If, for example, I reflect on my desire to become a nun, feel constrained by that desire, and strongly want to repudiate it,[80] then I am alienated from it and it is not autonomous. Perhaps, though, I do not feel alienated until I remember that I was forced to repeatedly watch *Sister Act* as a child. If I would repudiate the desire to be a nun only on remembering those experiences, then I am alienated in light of how the desire was formed, so it is not autonomous. Importantly, in Christman's model, there is no need for the person to actually engage in critical reflection. It is enough that if they did reflect, then they would not be alienated.[81] Furthermore, this hypothetical reflection need not be purely rational. It includes emotional elements;[82] and does not require someone to radically question everything about themselves, instead working against the background of other elements of their practical identities.[83] Even this hypothetical reflection, however, must not be the outcome of 'constriction, pathology, or manipulation'.[84] If I would desire to become a nun if Mother Superior would otherwise threaten my family's lives, then that is not enough to make the desire autonomous.

Non-alienation is a distinctive feature of Christman's theory. It is more usual for dual theories to require that a person 'identify' with or 'endorse' a desire or value on reflection.[85] Christman's motivation for this change is realism. He wants to recognise the widespread ambivalence that people feel about their own desires without necessarily saying that those people are not autonomous.[86] Requiring someone identify with or endorse a desire does not leave room for ambivalence; but requiring only that someone is not alienated, and does not

[79] *Ibid* 146, 155.
[80] *Ibid* 143.
[81] *Ibid* 155.
[82] *Ibid* 118, 144.
[83] *Ibid* 123-127, 131.
[84] *Ibid* 146-147, 155.
[85] For example Frankfurt (n 15), Dworkin (n 13).
[86] Christman (n 18) 143.

'reject' or 'repudiate' the desire,[87] allows for the more complex mixed feelings people often do have.[88] Because this aspect of Christman's model is relatively unique, however, footnotes in the following two sections do note those places where an identification-based theory would make minor differences.

For Christman, only those who meet both the competence and authenticity conditions are autonomous,[89] and autonomy can 'pick out' who should be treated paternalistically.[90] The next two sections contest this practical claim. They examine Christman's responses, and the possible alternatives, to two questions: what standard of reflection people should be held to and how should it be assessed? They show that none of the available options can justify actions against a person's expressed wishes without either relying on particular conceptions of the good or reaching conclusions, such as the idea that virtually nobody's expressed wishes should be respected, that the proponents of dual theories would almost certainly repudiate.

5. What standard of reflection is required?

If a test of autonomy includes a lack of alienation on critical reflection, then anyone applying that test needs to know what standard of reflection is demanded and how it is to be assessed. This section addresses standards of reflection and the next one addresses methods of assessment. Standards of reflection fall into four categories, all of which are canvassed here. Only one of them, 'bare reflection over time' is both neutral with regard to

[87] *Ibid* 144.

[88] A problem 'non-alienation' does not avoid is ambiguity. Christman appears to oscillate between alienation as a literal failure to recognise some part of the self as part of self (*ibid* 84, 148) and alienation as merely intense feelings of rejection towards some aspect of the self that is known to be part of the self (101, 143, 144, 146, 155). Although the latter interpretation seems more faithful to Christman's overall model, it nevertheless leads to difficult questions in practice. It treats alienated characteristics as not part of the person's deeper self, but the person will believe that those characteristics *are* part of their self; and, indeed, may only reject the characteristic *because* they believe that it is part of themselves. An addict attempting to recover repudiates their own addiction far more intensely than they repudiate the addiction of a stranger. For reasons of space, this issue is not dealt with in detail here.

[89] *Ibid* 15, 162.

[90] *Ibid*. He is not entirely consistent about this. At one point, he suggests that it is acceptable to act paternalistically towards those who lack competence, but not towards competent people who are alienated from a relevant motivation (162). Nevertheless, he seems to hold the stronger position more consistently (15, 110-111, 157, 173). Furthermore, as Killmister points out, he consistently holds that political respect is only owed to autonomous agents, and that competent-but-alienated agents are not autonomous; so a political obligation to treat competent-but-alienated agents non-paternalistically does not follow from his theory as a whole. Suzy Killmister, 'Autonomy, Liberalism, and Anti-Perfectionism' (2013) 19 Res Publica 353, 360. In any case, subsection 6.2 shows that a 'mere competence' test also fails to justify acts against a person's expressed wishes.

different conceptions of the good and can be applied without either producing unpalatable consequences or simply rendering all actions against a person's expressed wishes illegitimate. Section six then shows that combining this one remaining standard, bare reflection over time, with attention to how reflection is assessed shows that there are no options that can legitimate actions against a person's expressed wishes without either relying on norms that are not justified from the person's own perspective or producing unpalatable results.

5.1 Accurate self-evaluation

A superficially appealing standard of reflection is accurate self-evaluation: the person should be able to accurately assess both their own desire and their response to it. Christman disavows this option.[91] As he notes, there is good empirical evidence that most people, most of the time, are not capable of realistic self-assessment.[92] Indeed, the evidence makes the point even more strongly than Christman does. Self-assessments of broad traits like 'intelligence' or 'maturity' are systematically distorted,[93] and 'alienation' may be a broad trait: Christman himself characterises it impressionistically as 'resistance', 'repulsion', 'rejection', and 'repudiation'.[94]

The problems with a realistic self-assessment standard run deeper than this. One of the very factors that makes the concept of autonomy appealing, the insight that people can access their own emotions and 'inner' states in a way that other observers cannot, actually undermines the idea that people are capable of accurate self-assessment. No-one can experience other people's lives from the inside, so a state of 'pluralistic ignorance' prevails: most people believe that they are more self-conscious, ambivalent, and inhibited than average, and think that they feel emotions more strongly than most.[95] Obviously, the majority of people cannot be above average, so human self-evaluation must be systematically inaccurate, and it is probably inaccurate *because* we have privileged insight

[91] Christman (n 18) 127-128. He is not entirely consistent. Friedman observes that he sometimes relies on accurate self-evaluation when responding to critics. Marilyn Friedman, 'Review: Christman, John. The Politics of Persons: Individual Autonomy and Socio-Historical Selves' (2011) 121 Ethics 424, 428 (discussing Christman (n 18) 161-162).
[92] *Ibid* 128-129. See also David Dunning, *Self-Insight: Roadblocks and Detours on the Path to Knowing Thyself* (Taylor and Francis 2005); Lisa Bortolotti, *Irrationality* (Polity Press 2015).
[93] Dunning, *ibid* ch 6.
[94] Christman (n 18) 101, 143, 144, 146, 155.
[95] Dunning (n 92) 93-98.

only into our own minds. We have no direct comparator to measure them against.[96] Significantly, this pluralistic ignorance may actually drive feelings of alienation. Some evidence suggests that those who wrongly believe they are 'different' to others feel less commitment to shared projects.[97] In other words, far from being able to impartially assess their own alienation, people may feel alienated largely because they cannot impartially assess themselves at all. Furthermore, there are other barriers to accurate self-assessment. For instance, memories are not the neutral retrieval of objective facts, but tend to be systematically rewritten to 'preserve a coherent self-image'.[98] If autonomy requires non-alienation on reflection, and reflection must be historically accurate,[99] then this presents a challenge. A person's memories of how their desires were formed come to them already systematically distorted to preserve their self-image, which is to say to reduce alienation.

If accurate self-assessment is, at best, rare, then it provides unpromising grounds for neutral assessments of a person's autonomy. If a person's wishes only have to be respected when they accurately self-evaluate, and people generally cannot accurately self-evaluate, then very few wishes would ever demand respect. A system like this would be politically unpalatable, but it would also demand further justification. Dual theorists usually take themselves to be articulating the intuition that a person's wishes usually demand respect,[100] but an accurate self-evaluation standard would have exactly the opposite effect. It would instead justify a situation in which only the wishes of, at most, a few exceptional individuals had to be respected. Such wholesale legal change has not been argued for by dual theorists, nor would many support it; but, furthermore, justifying it would require an appeal to norms beyond the dual model itself. Anyone attempting to justify such a change would have to provide reasons for holding people to a standard of self-insight that most will not achieve.[101] Such reasons are not provided by dual models themselves.

[96] *Ibid.*
[97] *Ibid* 96.
[98] Bortolotti (n 92) 136.
[99] Christman (n 18) 155.
[100] For instance: *ibid* 13-14.
[101] One fall-back position is that 'accuracy' is not required, but 'rationality' is. This goes nowhere, even on relatively minimal definitions of rationality. As Bortolotti says, 'some of the deviations from good reasoning and clear thinking that are considered to be paradigmatic of irrationality are so widespread in the non-clinical population as to count as the statistical norm' (n 92) 49. See also subsection 5.3.

5.2 Bare reflection

Bare reflection provides an alternative standard to accurate self-evaluation. On this standard, 'reflection' would merely mean that someone is aware of their own relevant desire; and can form a response, whether acceptance or alienation, to that desire. The standard is easily achieved. If someone can intelligibly respond to another person asking them 'what do you want?' and 'are you sure?', then they have met it. This standard has influenced Christman. Although he does not adopt it, he recognises that it can help to fill the justificatory gap left by abandoning the idea that people can accurately self-evaluate. For this reason, he grounds his justification in Atkins's view that selfhood has an irreducible first-personal element.[102] In other words, he starts from the premise that what it is about each person's point of view that makes it 'theirs' is not reducible to anything else; and that this irreducible subjectivity demands respect.[103] In this way, bare reflection provides some content to the common intuition that I know myself better than others know me.

A bare reflection standard, unlike accurate self-evaluation, does not lead to the conclusion that very few people are ever autonomous. It is not, however, capable of justifying actions for a person's benefit against their expressed wishes. An example illustrates the problem. If I reach for a plate of chips against the better interests of my cardiac health, I might not meet a bare reflection standard. Perhaps I grab chips automatically, unaware even of my desire for them. This lack of reflection, however, is not enough to justify someone else stopping me from eating the chips. If I am merely unaware of my desire for chips, then I am not necessarily alienated from it, so it is not necessarily non-autonomous.[104] At most, my lack of reflection justifies a third party prompting it, perhaps by asking 'chips for breakfast again, Paul?' in the right tone. The real problem for a bare reflection standard, however, is

---

[102] Christman (n 18) 83, citing Kim Atkins, 'Narrative Identity, Practical Identity and Ethical Subjectivity' (2004) 37 Continental Philosophy Review 341. Similarly, he accepts James's distinction between the 'Me-self' (those characteristics of the self that a person can reflect on) and the 'I-self' (the mechanism of subjectivity that makes reflection possible): 30-31, citing William James, *The Principles of Psychology* (Henry Holt and Company 1890) 371.

[103] Kim Atkins, 'Autonomy and the Subjective Character of Experience' (2000) 17 Journal of Applied Philosophy 71.

[104] Dual theories that rely on 'identification' rather than 'non-alienation' fall into a different problem on a bare reflection standard. If someone has not reflected, they will have not 'identified' with the desire, so it is heteronomous and an action against their wishes is permitted. Most actions are unreflective even on bare reflection, so therefore most people's immediate wishes would not demand respect. This is the concern that led to Christman proposing a non-alienation standard. If an identification-based dual model tests merely whether the person *can* reflect, then the test will no longer reliably track dual-model autonomy and so cannot be justified by an appeal to it, see subsection 6.2.

with my potential responses to this prompt. When the possibilities are examined, none can justify acting against my will. If I say 'chips are better than a long life', then I am not alienated. If I stop reaching for the chips, then no action against my wishes is necessary because the process of reflection has changed my desire. Even, however, if I say 'I wish I didn't like chips so much' but continue to reach for the plate, some reason is needed for favouring my words over my actions. Both my words and my actions follow from my reflection. It is possible to describe my words as a second-order reflective response to first-order actions; but it is also possible to describe my actions as a third-order reflective decision that I am going to eat the chips despite my second-order qualms.[105] There may be ways through this regress, but a bare reflection standard does not supply them. The final possibility is that I am so irrationally consumed by my love of chips that I am unaware of the prompt. Even then, however, no action against my wishes is justified. The prompter must distinguish between two situations: one in which I have reflectively decided to ignore them and feast on chips, and one in which I am genuinely unreflective. The simple fact that I am ignoring them does not help them to make this distinction. It would be true in either case.[106] They may, however, try to draw on contextual evidence. Perhaps, for instance, I have been saying 'I really hope I don't end up eating breakfast chips again' all week, so they are tempted to say that I am not now reflective. This reasoning is not as innocent as it appears. It effectively adds a requirement that my reflection be consistent. This may be justifiable, but it is no longer a bare reflection standard.[107] Because a bare reflection standard puts no constraints on what counts as reflection, it cannot distinguish between a reflective refusal to engage and a lack of reflection. Consequently, it cannot justify actions against a persons expressed wishes.

5.3 External standards

If accurate self-evaluation sets the standard unrealistically high but bare reflection cannot legitimate actions against a person's expressed wishes, an obvious temptation is to seek

[105] This is a general and fundamental problem with dual theories: Gary Watson, 'Free agency' in Gary Watson (ed) *Free Will* (Oxford University Press 1982) 96, 108.

[106] If it is objected that the test should be of hypothetical reflection, not actual reflection, then a bare reflection standard faces an exaggerated form of the problems discussed in subsection 6.3, combined with the additional implausibility of claiming to be able to decide what the outcome of a person's hypothetical reflection would be at the same time as being unable to determine whether the actual person has even reflected.

[107] It is probably a 'bare reflection over time' standard, see subsection 5.4.

some standard between the two extremes. The problem with this strategy is that to be useful in practice the line between the two extremes must be drawn in some in definite place; and, wherever that line is drawn, it must be justified. In almost all cases that justification will draw on standards external to the dual model itself, and those standards will rely on particular conceptions of the good.[108] In other words, this strategy amounts to an effective abandonment of the idea that whether or not a person is autonomous can be assessed in a value-neutral way. Some dual models expressly deny that they can be value-neutral, but these models are not relevant here.[109] They do not claim to be able to underwrite value-neutral assessment anyway. More relevant to this chapter is the way that even models that do claim neutrality about the content of a person's desires still implicitly rely on external standards. Christman provides an example. He requires that reflection 'is not constrained by reflection-distorting factors' that he treats as factual, but which are clearly normative.[110] In particular, he asserts that we 'independently know' that 'constriction, pathology, or manipulation' 'effectively prevent minimal self-understanding'.[111] He provides almost no detail here. In particular, he does not examine how we 'independently know' this. He presents this 'knowledge' as a simple question of fact. Constriction, pathology, and manipulation, however, are all contested normative concepts that draw on particular conceptions of the good for persons.

Pathology may seem to be the most objective of Christman's three terms, so it is worth addressing first. He deals with the matter so shortly that it is difficult to be sure,[112] but his underlying view seems to be the 'medical model' view that psychiatric diagnosis consists of the value-free classification of symptoms and disorders.[113] This is unjustifiably naïve. In reality, diagnosis always includes a normative component, centred as it is on the idea of *dys*function.[114] Indeed, there are no particular mental experiences that, in themselves, are

[108] Freyenhagen and O'Shea (n 70) 55 point out that there is substantial pluralism even about 'epistemic norms' such as 'the degree of warrant required of beliefs that inform the deliberative process'. In other words, even deciding what should count as good evidence about the world requires drawing on norms that the assessed person may not share. Bortolotti (n 92) 147 makes the same point about rationality: it is itself 'a value concept' that not all people share the same version of, or value to the same degree.
[109] They are discussed in chapter five, subsection 2.1.
[110] Christman (n 18) 155.
[111] *Ibid* 147.
[112] *Ibid.*
[113] KWM Fulford and others, *Oxford Textbook of Philosophy and Psychiatry* (Oxford University Press 2006) 565.
[114] *Ibid* 570-573.

diagnostic of a psychiatric disorder: even experiences that would popularly be associated with delusions occur 'in both normal and pathological forms'.[115] As Freyenhagen and O'shea note, the diagnostic manuals are littered with 'normatively charged criteria': words like 'undue', 'excessive', and 'peculiar' that rely on some conception of what is appropriate for persons.[116]

Psychiatric assessment is value-laden; but, beyond that, the particular values underlying psychiatric assessment are often difficult to analyse, internally conflicted, and contested.[117] To point this out is not necessarily to criticise psychiatry. It is something many in the discipline recognise,[118] and it may be navigating this complex mix of facts and values better than any plausible alternative arrangement could. If, however, psychiatric diagnosis is inescapably value-laden, then 'pathology', left unanalysed, cannot be treated as a question of fact. Indeed, to hope that an unchecked concept of pathology can be safely used when determining whether actions against a person's expressed wishes are permissible is to wilfully ignore the lessons of history: for example, the political abuses of psychiatry that occurred in the twentieth century USSR.[119]

The same point applies to coercion. It is, as Raz says, 'an evaluative term'.[120] This point is even true of coercion by threats. To threaten someone is to communicate to them that if they act in particular way, then you will bring about some consequence they fear.[121] This 'empirical' definition, however, also captures a huge swathe of behaviour that is not usually treated as coercive. For instance, it might include telling a student that unless they include some references in their essay, you will not give them a passing grade. Getting a passing grade does motivate many students to include references, but it would be unusual to declare their academic work non-autonomous by reason of coercion. For something to count as coercive requires an additional evaluative judgment about 'how evil the threatened consequence' is.[122] Furthermore, some things that do not meet the empirical

[115] *Ibid* 569.
[116] Freyenhagen and O'Shea (n 70) 58.
[117] Fulford (n 113) 598.
[118] *Ibid* ch 21-22. This is a standard text.
[119] KWM Fulford and others, 'Concepts of Disease and the Abuse of Psychiatry in the USSR' (1993) 162 British Journal of Psychiatry 801.
[120] Joseph Raz, *The Morality of Freedom* (Clarendon Press 1986) 148.
[121] This is a loose approximation of *ibid* 149, A(1)-(4).
[122] *Ibid* 149.

criteria of a threat are nevertheless considered coercive. English law, for example, recognises that another form of 'unacceptable conduct …arises out of a relationship between two persons where one has acquired over another a measure of influence, or ascendancy, of which the ascendant person then takes unfair advantage'.[123] 'Unfair' in this context is clearly an evaluative term; but, beyond that, the courts have deliberately refused to give any 'comprehensive' test for this form of undue influence.[124] As with 'pathology', 'coercion' is not, as Christman presents it, a neutral question of fact. It is a complex mixture of fact and value.

'Constriction', the last of Christman's three terms, follows the same pattern. Christman's example of reflection-impairing constriction is 'having been denied minimal education and exposure to alternatives'.[125] Even this is not a question of fact. It must rely on some judgement about what constitutes 'minimal' education, and that judgement must rely on complex value judgements. What counts as a minimal education for a girl will not be the same in Manchester and in Taliban controlled Helmand province. Furthermore, even what counts as 'education' is not a question of fact. If my travelling family kept me entirely out of school but exposed me to many different places, taught me how to repair any vehicle, and imparted excellent business skills, then have I been denied education or merely given an alternative form of education? Any answer to the question must rely on value judgments. As with pathology and coercion, 'constriction' is not value-neutral.

Lurking behind all three of these normative terms there may be another, inchoate, idea that Christman would almost certainly not explicitly endorse. That other idea is 'norm' in the statistical sense of what is typical for a particular population. In the terms found in the diagnostic manuals, such as 'excessive' or 'peculiar', there seems to be an implicit comparison to the rest of the population. Indeed, for the word 'peculiar' to even have meaning, there must be some sense of what is 'usual'. Similarly, judicial reluctance to define 'undue influence' means that whether a particular influence is undue can only by determined by comparing it to what is normal in relationships of that particular type in this society. Constriction is a less clear concept, but the example Christman uses, of being denied 'minimal education', relies on the wider norms of society. In all three cases, there is a

[123] *Royal Bank of Scotland v Etridge (No 2)* [2001] UKHL 44, [2002] 2 AC 773 [8].
[124] *Ibid* [11].
[125] Christman (n 18) 147.

tendency to find the threshold, perhaps with a certain amount of leeway, in how people usually are and what people usually do. If people are to be held to such a standard, then it can only be because there is an unstated premise: 'it is always good to be as people usually are'. This is unpromising ground for any theory of autonomy. If implemented, then society would effectively say to the individual 'we will respect your ability to be self-determining, but you are only self-determining insofar as you are the same as everyone else'. No strand of theory spun from Mill's defence of individuality deserves this fate.

If autonomy relies on concepts such as pathology, coercion, and constriction; then, for autonomy to be neutral between different conceptions of the good, these ideas must be capable of working in a way that is similarly neutral.[126] They do not operate in this way at the moment; so to rely, as Christman does, on these ideas without further analysis creates a gap in justification. Any practitioner facing, for instance, the question 'does this person lack autonomy due to coercion?' cannot answer it without deciding whether the particular circumstances in question count as coercion. If they cannot answer that question in a value-neutral manner, then they cannot resolve the question of autonomy in a value-neutral manner. Life, unlike theory, presses the hard questions. Despite this weakness, however, Christman's model does contain another standard: one that appears to be able to legitimate actions against a person's expressed wishes in a value-neutral way.

5.4 Bare reflection over time

A final standard of reflection in Christman's work promises to evade the problems with accurate self-assessment, bare reflection, and the unacknowledged use of external standards. For him, reflection must be 'sustained …over a variety of conditions in light of the historical processes' that gave rise to whatever characteristic is being reflected on.[127] These criteria can be broken into two parts. 'In light of the historical processes' reduces to another unacknowledged reliance on norms from outside the model.[128] The 'sustained' and

[126] If it is argued that these concepts are only relevant insofar as they 'impair reflection' and that reflection is a question of fact, then the position collapses into circularity. What is the standard of reflection to which a person is being held, so that their reflection can be judged 'impaired'?

[127] Christman (n 18) 155.

[128] Christman stipulates that the historical processes reflected on should be 'adequately described', and that the process is adequately described when it is 'consistent with accepted evidence and known causal sequences' (154). If 'consistent with the accepted evidence' means logically consistent, it is imposing a standard that few people actually meet (Bortolotti (n 92) 127-136) and faces the same problems as accurate self-evaluation. If it means the evidence that should be accepted according some other standard, then it is another disguised appeal to poorly specified norms from outside the model. Similarly, 'known causal

'over a variety of conditions' conditions are more interesting for current purposes. They offer a standard of reflection that can sometimes legitimate actions against a person's expressed wishes without relying on unanalysed concepts from beyond the model. Imagine Adrian meets his abusive ex-wife, Helen, when shopping. She insists he goes back to her flat to 'catch up on old times'. Adrian is fearful, but he reflects for a moment and agrees. On a bare reflection standard, Adrian's decision is autonomous. It is quite possible, however, that he has also reflected on his desire to please Helen many times when she was not there, and perhaps he strongly repudiated that desire. In other words, he had 'sustained' alienation 'over a variety of conditions'. This suggests a 'bare reflection over time' standard.[129] Adrian's previous reflections are sustained over a variety of conditions, but his current desire to please Helen is not. A distinction can be drawn without relying on conceptions of the good that are not Adrian's own.

The freedom to act against someone's expressed wishes granted by a bare reflection over time standard is extremely limited. Imagine a slight variation in the facts. On this version, instead of consistently rejecting his desire to please Helen, Adrian sometimes felt alienated; but sometimes wholeheartedly endorsed all of the times that she said he was 'being too sensitive for a man'. In these circumstances, a bare reflection over time standard cannot distinguish between Adrian's occasional endorsement of his desire to please Helen and his occasional rejection of that desire without relying on some further justification. Both his reflective alienation from and his reflective endorsement of his desire are sustained over a variety of conditions. If one is to be favoured over the other, it must be for some reason other than meeting the procedural criteria. In other words, bare reflection over time can only justify actions against a person's expressed wishes without relying on external standards when two conditions exist: (a) the person would be alienated from the wishes on sustained bare reflection over a variety of conditions; and (b) the person would not endorse the wishes on sustained bare reflection over a variety of conditions. Already, this is a considerable constraint. Beyond this, however, the next section shows that, in application, even this standard must draw on other values to justify actions against a person's expressed wishes. In other words, dual models do not provide an account that can justify acting

sequences' assumes an agreed, factual, account of mental causation. As with 'pathology', Christman here relies on science that simply does not exist (Fulford (n 113) ch 15).

[129] For the purposes of exposition, this passage puts to one side Christman's, already discussed, simultaneous reliance on concepts such as 'coercion'.

against a person's expressed wishes while remaining neutral between different conceptions of the good.

6. How should reflection be assessed?

A bare reflection over time standard seems to promise a standard for evaluating a person's autonomy without relying on other contested norms. Any assessor of another person's autonomy, however, will need more than a standard. They will also need some way of measuring the person against that standard. When it comes to reflection, there are three possibilities: assessing the person's actual reflection, merely assessing their competence to reflect, and assessing what their response would be if they did, hypothetically, reflect. None of the three can be combined with the bare reflection over time standard in a value-neutral way that gives a single, determinate, answer.

6.1 Real reflection

Although his model does not formally require it, Christman occasionally seems to envisage assessing a person's actual process of reflection. For instance, when discussing a particular example, he says 'she may not meet the requirement of adequate reflection needed for autonomy'.[130] This is a little odd. Formally, Christman's model requires hypothetical reflection,[131] so the adequacy of someone's actual reflection should be irrelevant. This particular ambiguity illustrates a broader structural issue. Depending on how the hypothetical reflection is determined, any model based on it can, in practice, collapse into one based on real reflection. As Nys says, if an assessor must determine the outcome of a person's hypothetical reflection by asking them to actually reflect, then the model simply requires real reflection.[132] This is not the only reason, however, for considering the assessment of real reflection. It has at times been directly endorsed by other authors, such as Bratman.[133]

[130] Christman (n 18) 176.
[131] *Ibid* 155.
[132] Thomas Nys, 'Review: The Politics of Persons: Individual Autonomy and Socio-Historical Selves, by John Christman' (2011) 19 European Journal of Philosophy 474, 478.
[133] Michael E Bratman, 'Identification, Decision, and Treating as a Reason' in *Faces of Intention: Selected Essays on Intention and Agency* (Cambridge University Press 1999) 185. Bratman's view has since changed, (n 37) 6.

Assessing real reflection is hard to justify. Most people have not reflected in a sustained way on many of their desires and values; so, as with the bare reflection standard,[134] there is no warrant for acting against their expressed wishes. If someone has not reflected, then they are not alienated on reflection; so, on a model based on non-alienation, they are autonomous.[135] At most, all that is justified is prompting the person to reflect, but once reflection has been prompted there will be no justification for actions against the person's wishes. Once again, if I say 'I wish I didn't love chips for breakfast' after reflection, but eat them anyway, reasons are needed for favouring my verbal expression of alienation over my physical expression of endorsement. If someone has not previously reflected, then testing real reflection cannot justify actions against their expressed wishes.

In contrast, if someone has previously reflected, then testing real reflection may seem to fare better. In the example in the last section, Adrian's sustained alienation when reflecting on Helen's abuse did plausibly prevail over his fleeting endorsement of his desire to please her. The wider picture, however, is not so straightforward. Even when someone has engaged in sustained reflection, relying on real reflection does not always yield convincing answers. Adrian's sustained reflection was highly specific. He was reflecting on a particular situation that he had been in many times before. Life is not always like that. Often, we know the details of future events only roughly, but we know our past and present in great depth. Imagine, instead, that Adrian had reflected on what it would be like to be in an abusive relationship, but had never been in one. He then finds himself in an abusive relationship with Helen; and, reflecting on his situation, decides that she is not 'abusive', she is 'dominant'.[136] He reflectively endorses his desire to please her. In these circumstances, it is less clear why Adrian's previous sustained reflection should prevail over his current brief reflection. He can, with some justification, argue that his previous sustained reflection was naïve, and that he did not really 'know what it was like', but that his current fleeting reflection takes into account all the details of the situation: for instance, his feelings of

[134] Subsection 5.2.

[135] As with 'bare reflection' (subsection 5.2) models that rely on 'identification' or 'endorsement' instead of non-alienation will, for *real* bare reflection over time, conclude that most people are almost always heteronomous and that actions against their will are usually permissible. Very few, if any, dual theorists actually argue for this outcome and this problem seems to have motivated Bratman's move to hypothetical reflection (n 37) 6.

[136] If it is objected that he is wrong, and Helen *is* abusive, then his reflection is being held to either an accurate self-evaluation standard or an external standard. These face the problems discussed in, respectively, subsections 5.1 and 5.3.

loyalty towards Helen. Furthermore, because he remembers his previous reflection, his current reflection contains the older; but his older reflection does not contain the current. In these circumstances, it is still possible to argue that Adrian's previous reflection should be favoured, but it is necessary to argue for it. In other words, the model itself does not straightforwardly provide justification. It must draw on external norms, and if the assessment is to be neutral between different conceptions of the good, then those norms must also be neutral. As the discussion of external standards in subsection 5.3 shows, there is little evidence such standards exist.

Examples can be constructed to make real reflection seem workable, but it faces two hurdles when combined with bare reflection over time. If someone has not previously reflected, then it cannot justify actions against their expressed wishes. If someone has previously reflected, then it favours prior naïve reflection over current informed reflection in a counter-intuitive manner. If these counter-intuitive results are accepted, then a reason for accepting them seems necessary. If they are not accepted, then some way is needed for distinguishing between cases in which sustained reflection prevails and cases in which it does not. Both acceptance and denial therefore require appeals to external norms. Bare reflection over time simply cannot be combined with real reflection to legitimate actions against a person's expressed wishes without further justification.

6.2 Mere competence

One possible response to the justificatory problems that real reflection faces is to say that what is important is not the outcome of a person's reflection; but, instead, their mere ability to reflect. At one point, Christman seems to endorse this approach;[137] but, as Killmister remarks, nothing in his theory seems to support that endorsement, making it difficult to evaluate.[138] Nevertheless, other theorists, such as Dworkin, have more consistently based their dual theories on the ability to reflect, rather than real or hypothetical reflection.[139] Sticking with Christman's language, this sort of assessment can be called mere competence:

[137] Christman (n 18) 162.
[138] Killmister (n 90) 360.
[139] Dworkin (n 13) 15-16.

the person must be able to form intentions and reflect on them, but need not actually reflect.[140]

Autonomy in dual models requires non-alienation, or alternatively endorsement, on reflection; so requiring mere competence is not requiring the person actually be autonomous.[141] There is something very strange about this position. As Raz says, 'it is hard to conceive of an argument that possession of a capacity is valuable even though its exercise is devoid of value'.[142] It also leads to a gap in justification that can be shown by example. Imagine that Elena contracts a virulent strain of influenza that is likely to progress to septic shock over the next few days. Her doctor advises her that she may become delirious, and that a likely side-effect of delirium is that she will refuse treatment. Elena reflects and strongly repudiates this future refusal of IV antibiotics. In other words, she is alienated from that refusal. Here, a mere competence test produces bizarre results. Elena was alienated from her delirious refusal of treatment on sustained reflection; yet, when the time comes, no action against that refusal is permitted. She is *competent over time*, and that is all that the test demands. If it is objected that she accepted the treatment on reflection but we are respecting her refusal of it, then attention is being drawn to the actual content of the reflection, not the mere ability to reflect. In other words, a test of real reflection is being quietly substituted for one of mere competence.[143] If, instead, it is objected that she is not competent when delirious, then her reflection is being measured against a bare reflection standard,[144] and the 'over time' element has been quietly dropped. If, in a move few people would endorse in practice, she is left delirious long enough for the 'over time' element to be satisfied, then there is every chance that her refusal of treatment will be found competent. As discussed in subsection 5.2, 'bare reflection' merely means that someone is aware of their own relevant desire, and can accept or reject that desire. This is a standard almost all delirious people who are able to express a desire will meet; so this would lead again to a problem discussed in subsection 5.4: there are two competing sets of reflections,

[140] Christman (n 18) 155.
[141] Dworkin calls this competence 'autonomy' but immediately distinguishes it and the 'exercise' of autonomy (n 13) 17. This variation in terminology does not avoid the substantive problems discussed in this subsection.
[142] (n 119) 372.
[143] See 6.1.
[144] See 5.2.

one endorsing and one rejecting treatment, and no value-neutral way to choose between them.

A mere competence test faces further difficulties. When someone cannot reflect, it can authorise actions that the person would be unlikely to endorse if they were competent. Imagine Max is in the same situation as Elena, but is so severely needle-phobic that he cannot think about his future refusal of treatment at all. On a mere competence test, he can be given treatment despite his delirious refusal. He, unlike Elena, is not competent over time. This would be the case, however, even if he was a devout member of a religion that refused all medical treatment. What conclusion Max would be likely to reach if he could reflect is completely irrelevant to a mere competence test. In other words, the situation is possible in which Elena competently repudiates her refusal of treatment, yet that refusal of treatment must be accepted; but Max would almost certainly not repudiate his refusal of treatment, yet that refusal of treatment need not be respected. Testing mere competence produces results that have no necessary connection to dual model autonomy; so if such a test is justifiable, which seems unlikely, it must be due to some norm other than autonomy itself.

6.3 Hypothetical reflection

Christman largely avoids the problems associated with real reflection and mere competence. He does not require someone actually reflect to be autonomous, merely that they would not be alienated if they did so.[145] His motivation is realism. He acknowledges that many elements of everyone's personality were 'never chosen or reflectively appraised', but believes that many people are nevertheless autonomous.[146] Hypothetical reflection is a way of resolving these apparently contradictory commitments. For Christman, alienation can be tested in a way that gives assessments of the following kind: 'Shabab has not critically reflected on her desire to be a barrister; but if she did, then she would not be alienated'. The first part of this sentence, 'Shabab has not critically reflected', is worth pausing over. It shows the strangeness of Christman's position. His commitment to realism leads to him justifying assessments based on a fiction. It is, however, the clause to the right of the semicolon that contains the real difficulty.

[145] Christman (n 18) 145, 155.
[146] *Ibid* 145.

Statements of this type, 'if *P* did *x*, then *y* would happen', are called counterfactuals.[147] In this context, they lead to problems. A simple example, adapted from Wrigley, shows this: 'If I had woken earlier, I wouldn't have missed the 9am train'.[148] Imagine I woke at 8:55am, and it takes me 10 minutes to get to the station. In these circumstances, the statement may seem straightforwardly true. It is not. If I had woken up one minute earlier, then I might have still missed the train; but if I had woken up half an hour earlier, I might have caught it. This illustrates a difficulty with counterfactuals. They are indeterminate, so their audience usually defaults to a 'principle of charity' and interprets them in such a way that they are more likely to be true.[149] This, in itself, gives considerable power to the person who makes the counterfactual. It is tempting to respond to this problem by making the counterfactual more specific. This, however, is to misunderstand the nature of the problem. Counterfactuals are not indeterminate because they are overly general statements about this world, but because they are statements about what would happen in some imagined other world. Even the more precise statement, 'if I had woken up half an hour earlier, I wouldn't have missed the 9am train', is not straightforwardly true. It assumes, among countless other things, I would not have been fatally struck by a car while crossing the road to the station. In other words, it assumes that an alternative world in which I was not hit by a car is more similar to the real world than an alternative world in which I was hit by a car.

The train example might seem unproblematic. It might seem obvious which possibility is more similar to the real world. This appearance is, however, deceptive. For example, take John, who has recently died. John, when alive, never seemed to think about his own death, and he has left no funeral instructions. An argument has broken out between his daughters. Sally believes that if John had thought about it, then he would have wanted to be buried in the same plot as his parents; but Lucy says that John would want to be cremated and have his ashes scattered from the top of Pillar, a hill he loved. Sally points out that John always visited his parents' grave on his way to the pub on Sunday afternoons. Lucy counters that he always wanted his daughters to do more hill-walking; and that if his ashes are scattered from the hill, then they will. These assertions and counter-assertions are not simply factual,

[147] There is a large philosophical literature on counterfactuals. In particular see: David Lewis, *Counterfactuals* (first published 1973, 2nd ed, Wiley-Blackwell 2001).
[148] Anthony Wrigley, 'The Problem of Counterfactuals in Substituted Judgement Decision-Making' (2011) 28 Journal of Applied Philosophy 169, 179.
[149] *Ibid* 180.

even when they rely on true facts about John's behaviour and hopes. They are also claims about which facts count for similarity. Sally is not simply bringing a fact about John's behaviour to a neutral tribunal of similarity, and Lucy is not simply bringing a fact about his hopes. Sally is also claiming that what counts for similarity is John's behaviour, and Lucy is claiming that what counts is his hope. Lewis calls this a 'rule of accommodation',[150] by which counterfactual claims tend to validate themselves. There is no neutral way to arbitrate such conflicts because there is no single similarity-relationship that 'counts'. Instead, similarity is context-dependent.[151] Probably, Sally's suggestion is more similar to John's behaviour, and Lucy's suggestion is more similar to his hopes. If a decision is to be made, then any attempt to deal with this conflict by remaining in the land of counterfactuals goes nowhere. If Sally argues 'Dad was a creature of habit' so he would have chosen the graveyard, then she is still simultaneously making a both a claim about which world is more similar to the real world and a claim about what should count for similarity. If Lucy argues that 'Dad was a dreamer' so he would have chosen cremation, then she too is both claiming similarity and making a claim about what should count for similarity. Counterfactual similarity is normative all the way down.

When combined with a bare reflection over time standard, the indeterminate and normative nature of hypothetical reflection has catastrophic implications. As discussed in subsection 5.4, bare reflection over time can justify actions against the person's expressed wishes only if the person would be alienated from the relevant desire on sustained reflection over a variety of conditions *and* they would not authentically endorse it on sustained reflection over a variety of conditions. If a person's autonomy is to be neutrally assessed, then the test cannot simply be 'is a world in which P reflects and is alienated more similar to the real world than a world in which they reflects and are not alienated?' Counterfactual similarity is not normatively neutral. A stronger test, however, also fails. If the test is 'would P be alienated from, and not authentically endorse, the relevant desire in any imagined world in which they reflected?', then actions against a person's expressed wishes are never justified. No matter how out of character a desire might seem, it is always possible to construct a story in which the person is converted, like Paul on the Damascus road, to authentically endorsing it on reflection; and, on this version of the test, any such

[150] David Lewis, *On the Plurality of Worlds* (Blackwell 1986) 251-252.
[151] *Ibid* 254.

endorsement renders the desire autonomous. If it is argued that a particular conversion story is unlikely or even absurd, then these are evaluative judgments that rely implicitly on some idea of similarity. If, alternatively, it is argued that a particular counterfactual story does not fit the 'narrative' of a person's life, then this misses the central problem. Counterfactually, a person does not have one narrative of their life, but an ocean of possible narratives; and, as Lippitt points out, dual theorists offer no criteria for distinguishing one narrative from another.[152] Criteria could, of course, be supplied; but this, once again, is to draw on norms other than autonomy itself.

7. Conclusion

Legal scholars have drawn on dual models of autonomy when both justifying mental capacity law and when arguing for it to be changed, and the appeal of such models is obvious. They appear to replace the messy, value-laden question of when, if ever, it is appropriate to act against a person's expressed wishes with the putatively factual question of whether or not a particular expressed wish is attributable to the deep agent. This appearance is, however, a mirage: one that entirely dissolves when attention is paid to how a model could be used to evaluate a real human subject. Whoever made such an evaluation would need some standard of reflection to assess the person against. If, however, the standard is accurate self-evaluation, then virtually nobody will be found to be autonomous, a conclusion that no-one who relies on dual theories argues for. If, instead, only bare reflection is required, then autonomy can simply never justify acting against a person's expressed wishes. Again, this is not a position that any dual theorist argues for. Even those who advocate radical change under the banner of the UNCRPD distinguish between a person's 'verbal expression' and their 'true will'.[153] In the face of this dilemma, it is perhaps unsurprising that external norms are often smuggled into dual theories under the guise of 'independently known' facts, as with Christman's recourse to the concepts of pathology, coercion, and constraint. Despite these problems, however, one standard of reflection does seem to be able to distinguish between when it is and is not acceptable to act against a person's expressed wishes without drawing on further norms. A bare reflection over time

[152] John Lippitt, 'Getting the Story Straight: Kierkegaard, MacIntyre and Some Problems with Narrative' (2007) 50 Inquiry 34, 58.
[153] Flynn and Arstein-Kerslake (n 65) 98.

standard can justify actions against a person's expressed wishes if they would be alienated from those wishes on sustained reflection across a variety of conditions and would not endorse those wishes on sustained reflection across a variety of conditions.

Even if this neutral standard of evaluation exists in theory, however, there is no way to apply it in practice. Some way of measuring the person's reflection against the standard is required. If that method is to assess their real reflection, then actions against a person's wishes cannot be justified if they have not previously reflected on that wish. Furthermore, when the person has previously reflected, then testing real reflection sometimes yields counterintuitive results, favouring a sustained but naïve reflection over a brief but informed one. If, instead, a person's mere competence to reflect is assessed, then actions can be justified against an incompetent person's wishes even when they would almost certainly not be alienated on reflection, but not against a competent person's wishes even when they are factually alienated. Dual model autonomy is irrelevant to the results of this test, so it cannot justify it. Finally, hypothetical reflection, testing whether the person would be alienated if they did engage in sustained reflection, cannot distinguish between the vast numbers of hypotheticals that surround any actual facts without relying on a further, radically indeterminate, evaluative notion: 'similarity'.

As dual models do not isolate value-neutral criteria for determining when a person is autonomous, they cannot supply such criteria to law. This has implications for the three assumptions discussed in the introduction to this chapter. In particular, it is not possible to sustain both the idea that there is close association between 'autonomy' in law and 'autonomy' in political philosophy and the idea that whether or not a particular person is autonomous at a particular time can be assessed in a way that is neutral between different conceptions of the good. One or both of these ideas has to go. The next chapter examines the rhetoric of judges in the Court of Protection, and shows that even addressing this issue is a complex question. They do not talk about 'autonomy' consistently and what 'autonomy' is in mental capacity law cannot be straightforwardly read off the judge's speeches. It has to be reconstructed from their practices.

Chapter Three

The relationship between autonomy and adult mental capacity in the law of England and Wales

(A paper under review by Medical Law Review)

1. Introduction

This chapter is about a gap between judicial rhetoric and behaviour. Judges in particular cases describe the relationship between autonomy and mental capacity in a variety of ways, but if their decisions across the full range of cases are to be coherently accounted for, then the relationship cannot be as simple as they say it is. The existence of this gap does not imply that the judiciary have done anything wrong. Judgments are, as Harrington says, 'exercises aimed at persuading specific audiences of the truth of certain facts and the desirability of certain courses of conduct';[1] and published judgments are directed towards a variety of readerships with a variety of interests and needs.[2] In particular, those audiences, especially the parties, primarily interested in the outcome of the particular case can have different needs to those audiences more interested in the case's implications for the law as a whole. In mental capacity law, where the consequences of a decision are often profound and controversy is seldom far away, the needs of these two audiences can be in tension. The immediate audience need final resolution of an often prolonged conflict in a way that acknowledges that their concerns matter.[3] This demand is often met by simple, unequivocal declarations of principle: for instance, 'the purpose of the best interests test is to consider matters from the patient's point of view'.[4] There is nothing wrong with such declarations. They serve some purposes well. This chapter, however, shows that they do not always serve an audience attempting to understand the wider legal regime well. They focus attention too tightly on whatever issue is controversial in the particular case, neglect what is irrelevant

[1] John Harrington, *Towards a Rhetoric of Medical Law* (Taylor and Francis 2017) 2.

[2] For a discussion see EW Thomas, *The Judicial Process: Realism, Pragmatism, Practical Reasoning and Principles* (Cambridge University Press 2005) ch 10.

[3] Sir Mark Hedley, *The Modern Judge: Power, Responsibility and Society's Expectations* (LexisNexis 2016) ch 2.

[4] *Aintree University Hospitals NHS Foundation Trust v James* [2013] UKSC 67, [2013] 3 WLR 1299 [45] (Hale L).

to it, emphasise certainty over complexity; and so present a distorted view. When most published judgments, justifiably,[5] prioritise the immediate audience in this way, then extracting a coherent description of the legal regime from the patchwork of partial accounts can take some work.

One aspect of domestic mental capacity law exhibits such a patchwork of partial accounts. Judges tend to use whatever ideas about personal 'autonomy', a word that is not legally defined, best suit their rhetorical needs in the immediate case, whether or not their usage can coherently account for the law as a whole. As a result, three apparently contradictory accounts of the relationship between mental capacity and autonomy coexist. On one account, people with mental capacity about a matter are autonomous enough to make their own decisions about it, but those without such capacity are not. Mental capacity alone decides the issue of autonomy. On the second account, however, mental capacity alone is not enough for someone to be considered autonomous. Freedom from coercion and undue influences is also required. Finally, in the third account, even when someone lacks mental capacity about a matter they may be autonomous with regard to it. At first sight, these accounts appear to contradict one another; but all the cases in which they appear were decided within the current law, and it is possible to draw them together into a coherent structure that captures the important features of each. This unified account is given in section five.

This chapter evaluates a gap between how 'autonomy' appears if judicial comments in particular cases are taken at face value and how it appears if the law is treated as consistent across the full range of cases. It is not an attempt to evaluate practice against any already-determined philosophical account of autonomy, valuable as such work may also be.[6] Instead, it is an attempt to take seriously Wittgenstein's suggestion to start not with conceptual analysis, but in practice: to '*look and see*' how a word is used.[7] As he notes, 'looking' sometimes does not reveal a common core to a concept, one which would be present in all uses of it, but rather 'a complicated network' of overlapping similarities.[8] He

[5] After all, the function of the court is to decide the case. It 'should not be used as a general advice centre'. *R (Burke) v GMC* [2005] EWCA Civ 1003, [2006] QB 273 [21] (Lord Phillips MR).
[6] See, for instance, Camillia Kong, *Mental Capacity in Relationship: Decision-Making, Dialogue, and Autonomy* (Cambridge University Press 2017).
[7] Ludwig Wittgenstein, *Philosophical Investigations* (first published 1953, PMS Hacker and Joachim Schlute eds, GEM Anscombe and others tr, 4th ed, Wiley-Blackwell 2009) §66 (emphasis in original).
[8] *Ibid.*

goes on to suggest that when the uses of a concept lack a common core, then the scholarly assumption that a short, precise definition can capture the entire concept becomes hard to justify.[9] Nevertheless, it is often useful to start with a rough definition, if only to indicate a broad area of interest. For the purposes of getting started, then, autonomy is a person's capacity for self-government, in the sense that they are able to be in control of their own life.[10] This definition is broad enough to capture the judicial uses of the word that follow, but narrow enough to have an interesting feature. If autonomy is capacity for self-government, then it is not the same thing as the simple freedom from external restraints.[11]

In the courts, self-government and freedom from external restraints have an intimate, justificatory relationship,[12] which allows a typology of arguments that rely on autonomy to be made. There are two major types. First is the 'appeal to respect for autonomy':[13] if someone is believed to be capable of self-government, and self-government is believed to be good, then that is a reason to refrain from interfering in their life against their will. Such arguments are common in capacity cases, and can even determine their outcome.[14] Attention to this type of argument is central to the methodology here; for if it is to be successful, then that person must already be plausibly considered to be autonomous.[15] In other words, paying attention to when the appeal to respect for autonomy is successful offers a way to assess who is considered autonomous by the courts, one that that need not take individual judicial comments at face value. This appeal should, however, be distinguished from another argument, the 'appeal to the value of autonomy', which simply

[9] *Ibid* §69-71.

[10] Similar definitions occur in the courts. For example, Lord Sumption grounds autonomy in the 'moral instinct' that 'individuals are entitled to be the masters of their own fate': *R(Nicklinson) v MoJ* [2014] UKSC 38, [2014] 3 WLR 200 [208].

[11] Isaiah Berlin, 'Two Concept of Liberty' in *The Proper Study of Mankind: An Anthology of Essays* (first published 1958, Pimlico 1998) 191, 194; John Coggon and José Miola, 'Autonomy, Liberty, and Medical Decision-Making' (2011) 70 Cambridge Law Journal 523.

[12] This justificatory relationship also occurs in other contexts: For example: John Stuart Mill, *On Liberty* (Yale University Press 2003) 80-82 (and throughout).

[13] George Sher, *Beyond Neutrality: Perfectionism and Politics* (Cambridge University Press 1997) 16. See also Manne Sjöstrand and others, 'Paternalism in the Name of Autonomy' (2013) 38 Journal of Medicine and Philosophy 710.

[14] For example, *PC v City of York Council* [2013] EWCA Civ 478; [2014] 2 WLR 1 [53], discussed in section 2.

[15] Another way of framing this point might also be to ask who, in domestic law, has 'legal capacity', the ability to hold and exercise legal rights and duties: Convention on the Rights of Persons with Disabilities (adopted 13 December 2006, entered into force 3 May 2008) 2515 UNTS 3 (UNCRPD) art 12(2). 'Legal capacity' is, however, a term seldom seen in the domestic case law. In contrast, 'autonomy', and related concepts such as 'self-determination' and 'sovereignty of the individual', are extremely common.

asserts autonomy is valuable and should be promoted.[16] This latter argument can also determine the outcome of domestic cases,[17] but it operates differently. Basing a decision on the idea that autonomy is valuable and should be promoted does not require belief that the person is already autonomous. For that reason, although appeals to the value of autonomy are fascinating in their own right, they are not useful here.

In addition to excluding appeals to the value of autonomy, this chapter also takes the process of mental capacity assessment as largely unproblematic. This is not because assessments of capacity are an uncontentious matter. They are not.[18] Indeed, the Mental Capacity Act 2005 ('the MCA') directly ties incapacity to 'an impairment of, or a disturbance in the functioning of, the mind or brain';[19] and, although this is often broadly read, it is almost certainly a case of unjustified discrimination on the grounds of disability.[20] The focus in this chapter, however, is on what follows from a finding of incapacity, not on the finding itself; so these questions are largely put to one side. Similarly, none of the three accounts of the relationship between autonomy and capacity that appear in the courts directly address the Mental Health Act 1983 ('the MHA'). That Act undoubtedly affects whether or not a person will be considered autonomous by the courts, but mental capacity is a concept for the most part foreign to the MHA, so the subject is beyond the scope of this chapter. Finally, nothing here suggests that there is any legal conflict in the cases that are used as examples of the different accounts of the relationship between autonomy and capacity. Indeed, section five, which gives a coherent account, shows that there is not even any necessary conceptual conflict. The conflict is only in the rhetoric used, but the rhetoric used matters. It can lead to misunderstandings of the law. Mental capacity cases, rightly or wrongly, often decide the most important questions in a person's life: for instance, whether a woman will be allowed to live with her husband,[21] what contact a young man will have with his

[16] Sher (n 13) 16.
[17] For example, *Northamptonshire Healthcare NHS Foundation Trust v ML* [2014] EWCOP 2, [2014] COPLR 439 [44]-[45].
[18] UN Committee on the Rights of Persons with Disabilities, 'General Comment No 1' (2014) UN Doc CRPD/C/GC/1 para 27.
[19] MCA s2(1).
[20] Essex Autonomy Project, 'Achieving CRPD Compliance: Is the Mental Capacity Act of England and Wales Compatible with the UN Convention on the Rights of Persons with Disabilities? If Not, What Next?' (University of Essex 2014) 31-36.
[21] *PC* (n 14).

mother,[22] or whether a woman will be given a caesarean section against her wishes.[23] Clarity about both the law and its justification is therefore essential. With that in mind, it is best to start with the most oversimplified account.

2. The gatekeeper account

The least complex account of the relationship between autonomy and capacity is to treat mental capacity as the threshold for autonomy. On this view, if someone has mental capacity with regard to a particular decision, then they are autonomous with regard to that particular matter, so the state should not interfere with their decision. If, conversely, they do not have mental capacity with regard to the decision, then they are not autonomous, so the state need not exercise such restraint. As Donnelly says, capacity acts as autonomy's 'gatekeeper'.[24] Usually, gatekeeper accounts acknowledge that those without mental capacity have some capability for self-government. They only deny that it sufficient to be overriding in law. In other words, their distinguishing feature is not whether or not they think those without capacity have any capability for autonomy, but that they treat the presence or absence of mental capacity as the sole criteria for deciding whether an appeal to respect for autonomy will be successful.

2.1 Recent examples of the gatekeeper account

The gatekeeper account was dominant by the end of the last century. In *Bland,* Lord Mustill stated that 'if the patient is capable of making a decision …his choice must be obeyed';[25] but someone's 'inability to make' a necessary decision allows someone else to decide in that person's 'best interest'.[26] This idea was expressly linked to autonomy in *Re C (Adult: Refusal of Medical Treatment)*, 'if the patient's capacity to decide is unimpaired, autonomy weighs heavier, but the further capacity is reduced, the lighter autonomy weighs'.[27] In *Re MB*, Lady Justice Butler-Sloss made it clear that this is an exclusive test: 'the only situation in which it is lawful for the doctors to intervene is if it is believed that the adult patient lacks the

[22] *Re PL* [2015] EWCOP 44.
[23] *The NHS Acute Trust v C* [2016] EWCOP 17.
[24] Mary Donnelly, *Healthcare Decision-Making and the Law: Autonomy, Capacity, and the Limits of Liberalism* (Cambridge University Press 2010) 90.
[25] *Airedale NHS Trust v Bland* [1993] AC 789 (HL) 891.
[26] *Ibid* 892.
[27] [1994] 1 WLR 290 (Fam) 292 (Thorpe J).

capacity to decide'.[28] These cases all deal with medical treatment, but a Law Commission report from the same period suggests that the gatekeeper view already had broader application: 'a person with capacity is entitled to refuse community care or other protective services'.[29] That report, and the concurrent case law, eventually contributed to the definition of incapacity found in the Mental Capacity Act 2005.[30]

The MCA does not mention autonomy. Nevertheless, a gatekeeper account of the relationship between autonomy and capacity helped to justify the Act. On the second reading of the Mental Capacity Bill in the House of Lords, the Lord Chancellor stated that its first aim was to allow 'adults to take as many decisions as they can for themselves' but that 'where adults cannot make decisions …the Bill ensures that the decisions which are made for them are made in their best interests'.[31] By omission, this implies more than it states. In these passages, it seems as though the MCA applies to all adults who 'cannot make decisions', not only to those adults who cannot decide 'because of an impairment of, or a disturbance in the functioning of, the mind or brain'.[32] The possibility of people being unable to decide for other reasons is ignored, and this tends to imply a gatekeeper view of the relationship between autonomy and capacity: those with capacity are capable of self-determination, and those without are not. This became explicit in the Act's Code of Practice, which says that 'an adult …has full legal capacity to make decisions for themselves (the right to autonomy) unless it can be shown that they lack capacity'.[33] To some extent, the Act's test for incapacity reflects this gatekeeper view, for it treats mental capacity as an either/or concept. At the 'material time',[34] someone can either make a particular decision or they cannot, and the difference between these two conditions is considerable. The Act creates many powers affecting how decisions are made for those who lack capacity: a general defence for actions made in their best interests;[35] a section allowing the donee of a

[28] [1997] EWCA Civ 3093, [1997] 2 FCR 541, 555.
[29] Law Commission, *Mental Incapacity* (Law Com No 231, 1995) para 9.10.
[30] MCA s2(1): 'A person lacks capacity in relation to a matter if at the material 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'. 'Unable to make a decision' is further defined in s3(1) as an inability to understand, retain, or 'use or weigh' the information relevant to the decision, or an inability to communicate the decision.
[31] HL Deb 10 January 2005, vol 668, col 12 (Lord Falconer LC).
[32] MCA s2(1).
[33] Department for Constitutional Affairs, *The Mental Capacity Act 2005 Code of Practice* (The Stationery Office 2007) para 1.2.
[34] MCA s2(1).
[35] *Ibid* s5.

Lasting Power of Attorney to make such decisions;[36] a power of the court to make such decisions or to appoint a deputy to make them;[37] and a power for individuals with capacity to make an advance decision to refuse treatment at a future time when they lack capacity.[38] It also allows the court to make interim orders in someone's best interests where there is 'reason to believe' that they lack capacity in relation to a matter.[39] In contrast, the Act creates no powers at all to make decisions for those with capacity. This contrast lends the gatekeeper account a certain plausibility.

This plausibility makes it unsurprising that judges still sometimes give the gatekeeper account when making decisions under the Act. The account can be split into two limbs: (a) the idea that those with capacity in regard to a particular decision are autonomous, so the state should not interfere with what they decide; and (b) the idea that those without capacity in regard to a particular decision are not autonomous, so this restraint is not appropriate. Sometimes, the court states the whole gatekeeper account. For instance, in *Bailey v Warren*, decided between the MCA being passed and coming into force, Lady Justice Arden held that 'capacity is an important issue because it determines whether an individual will in law have autonomy over decision-making'.[40] More often, however, judges will expressly state only one of the limbs, while implying the other. For instance, in *PC v City of York Council,* Lord Justice McFarlane held that 'unless they lack mental capacity to make that judgment, it is against *their* better judgment …the statute respects their autonomy so to decide'.[41] Here, the emphasis is on the idea that a person with capacity is autonomous. Similarly, in *Nottinghamshire Healthcare NHS Trust v RC,* Mr Justice Mostyn found that 'even to think about imposing a blood transfusion' against the will of a detained patient with capacity, despite a power to do so existing under section 63 of the Mental Health Act 1983, would be 'an abuse of power'.[42] Although the word 'autonomy' does not appear in *RC*, the concept pervades it, with Mostyn J even quoting Mill on the sovereignty of the individual.[43]

---

[36] *Ibid* s9.
[37] *Ibid* s16.
[38] *Ibid* s24.
[39] *Ibid* s48.
[40] [2006] EWCA Civ 51, [2006] CP Rep 26 [105]. She continues, 'although no court should rush to interfere with the power of an individual who is competent to make decisions, it must not shirk the duty of intervening where an individual is not so capable and needs to be protected by the law'.
[41] (n 14) [53] (emphasis in original).
[42] [2014] EWCOP 1317, [2014] COPLR 468 [42].
[43] *Ibid* [8] citing Mill (n 12) 80.

These cases emphasise the autonomy of those with capacity, but others emphasise the lack of autonomy of those without capacity. For instance, in *Independent News and Media v A,* the Lord Chief Justice held that 'the responsibility of the Court of Protection arises just because the reduced capacity of the individual requires interference with his or her personal autonomy'.[44] This is a subtle sentence. It acknowledges that those without capacity may have some capability for self-government, while simultaneously stating that their lack of capacity is enough for appeals to respect for that autonomy to fail. Similarly, in *Re E (Medical Treatment: Anorexia)* it was found to be in a 32 year old woman's best interests to be fed against her will, despite this representing 'a very severe interference with her private life and personal autonomy'.[45] She was 'fully aware of her situation',[46] but because she was found to lack mental capacity, this did not restrain the court. In both of these cases, it is acknowledged that someone without capacity might have some ability to be self-determining. What is denied is that this ability is such that respect for it should restrain the state from acting against their wishes.

2.2 Function of the gatekeeper account

The gatekeeper account has been influential. It appeared in the appellate decisions that laid the groundwork for the MCA, and some recent examples also come from the Court of Appeal. The following sections show that it is, nevertheless, only a partial description of the law. This is not, however, to say that judges in these cases were in error. Rather, the account served important rhetorical purposes, and these are worth identifying. The two different types of case, those that emphasise the autonomy of a person with capacity and those that emphasise the lack of autonomy of the person without capacity, have two different purposes. When someone is making an unwise decision, but felt to have capacity, as in *PC* and *RC*,[47] then the gatekeeper account affirms that the matter is beyond the court's jurisdiction and so absolves it, and service providers, of any responsibility for the consequences. The ability to legally act on behalf of the person is treated as completely limited by the MCA.[48] For instance, in *PC*, 'we must leave PC free to make her own decision,

[44] [2010] EWCA Civ 343, [2010] 1 WLR 2262 [18] (Lord Judge CJ).
[45] [2012] EWCOP 1639, 2 FCR 523 [125] (Peter Jackson J).
[46] *Ibid* [5].
[47] In *PC,* the unwise decision was to live with her husband, who had a history of serious sexual offences. In *RC,* it was refusal of blood transfusions when they might be clinically indicated.
[48] In *RC,* Mostyn J does mention the inherent jurisdiction of the High Court, discussed in the next section; but neither applies it nor gives reasons why it cannot be applied (n 42) [13].

and hope that everything turns out well in the end'.[49] In contrast, in cases where the judge has concluded that is in someone's best interests to make a decision against their expressed wishes, the gatekeeper account affirms the court's authority to do so. It does so by positioning the court's determination as objective, and so decisive, but the person's wishes as merely subjective. For instance, in *Re E,* when determining to feed a woman against her will, 'E's life is precious, whatever her own view of it now is'.[50] Similarly, in *An NHS Trust v CS*,[51] Mr Justice Baker had to decide whether it was in a young woman's best interests to have an abortion. Prior to an assault by her partner that left her with significant brain damage, she intended to have the procedure; but following the assault, she wished to 'keep the baby'. Of these latter wishes, the judge found 'it seems to me impossible for this court to attach any significant weight to them bearing in mind her patent lack of capacity'.[52] As in *Re E*, the gatekeeper account helps to affirm the objectivity, and hence legitimacy, of the court's decision, by contrasting it with the subjectivity of the wishes of the person who is decided for.

The gatekeeper account, then, has two uses: absolving the court from responsibility when someone has capacity, and affirming its authority to act when someone does not. It is, however, a primitive tool. It is of little use when the autonomy of someone without capacity seems to demand respect or when someone with capacity seems, nevertheless, to be incapable of self-government. The latter of these situations is dealt with in the next section, for when someone with capacity seems to lack autonomy, then the judiciary may reach for the insufficiency account of the relationship.

3. The insufficiency account

In the gatekeeper account, if someone has capacity, then respect for their autonomy restrains the state from interfering with their wishes. The insufficiency account modifies this. Like the gatekeeper account, it holds that people without capacity do not have an overriding right to respect for their autonomy. Unlike the gatekeeper account, however, it does not hold that people with mental capacity necessarily do have an overriding right to respect for their autonomy. They may or may not have such a right, depending on other

[49] (n 14) [64] (Lewison LJ).
[50] (n 45) [122] (Peter Jackson J).
[51] [2016] EWCOP 10; [2016] COPLR 187.
[52] *Ibid* [21].

conditions. In other words, in the insufficiency account, mental capacity is necessary but not sufficient for autonomy to be respected as a right. Something else is also needed. As Lady Hale says in *R v Cooper*, 'Autonomy entails the freedom and the capacity to make a choice'.[53] This 'freedom', the thing beyond capacity that is needed for autonomy, is consistent in the cases that give this account. It is voluntariness, freedom from undue external influences.

3.1 Recent examples of the insufficiency account

The insufficiency account predates the Mental Capacity Act. In 1993, in *Re T (Adult: Refusal of Treatment),* Lord Donaldson found 'in some cases doctors will not only have to consider the capacity of the patient to refuse treatment, but also whether the refusal has been vitiated because it resulted not from the patient's will, but from the will of others'.[54] In other words, 'undue influence'[55] from a third party may have such an effect on someone with capacity that it will 'destroy her volition'.[56] When this happens, the court will not treat the affected person as so autonomous that their expressed wishes must be abided by.

Re T is older than the gatekeeper cases of *Re C* and *Re MB*, and those cases applied *Re T*, so it is worth examining how an insufficiency account became a gatekeeper one.[57] It did so by two routes, case law and Law Commission led reform. Taking the case law first, neither *Re C* nor *Re MB* concerned allegations of undue influence. As a result, when applying *Re T,* they did not explicitly mention that part of the precedent.[58] This omission led to a gatekeeper account: for instance, in *MB,* 'the *only situation* in which it is lawful for the doctors to intervene is if it is believed that the adult patient lacks the capacity to decide'.[59] Something similar happened with the Law Commission. *Re T* envisages two separate tests, first whether the person has capacity and then whether their will is nevertheless being

[53] [2009] UKHL 42, [2010] Crim LR 75 [27]. This criminal case was decided under the Sexual Offences Act 2003, but the quote is from a section that draws extensively on the MCA. 'Capacity' is effectively the same in 'civil and criminal jurisprudence': *R v A* [2014] EWCA Crim 299, [2014] 1 WLR 2469 [25] (Macur LJ). For largely pragmatic reasons, however, the test for capacity to consent to sex is, treated differently. As criminal cases concern a particular past event they assess the person's capacity to consent to sex with a particular other person; but as civil cases concern a person's future sexual encounters they assess capacity to consent to sex 'on a general and non-specific basis' *IM v LM* [2014] EWCA Civ 37, [2015] Fam 61 [77] (Leveson P).
[54] [1992] EWCA Civ 18, [1993] Fam 95 [37].
[55] *Ibid* [41] (Butler-Sloss LJ).
[56] *Ibid.*
[57] This is not to say that the gatekeeper account replaced the insufficiency account. The two coexist.
[58] *Re C* (n 27) 294; *Re MB* (n 28) 549.
[59] *Re MB* (n 28) 555 (Butler-Sloss LJ) (emphasis added).

overborne,[60] but the Law Commission's report considers that someone with a mental disability who is 'unable to exert their will against some stronger person who wishes to influence their decisions'[61] lacks mental capacity. Combining incapacity and involuntariness in this way necessarily leads to a gatekeeper account of the relationship between capacity and autonomy. Rather than being, as in the insufficiency account, something additional to incapacity that can remove a person's autonomy, involuntariness simply becomes part of incapacity itself. There is, however, a difference between this and the gatekeeper cases. The cases, by silence, seem to erase any consideration of involuntariness; but the Law Commission, by incorporation, draws involuntariness into the definition of incapacity itself.

The Law Commission's combination of incapacity and involuntariness has led to some confusion; for, despite the assertion that undue influence could go to incapacity, the actual wording of the Commission's draft Bill precludes this from happening in every case. The Bill states that someone lacks capacity if 'unable *by reason of* mental disability to make a decision'.[62] Similarly, the MCA, as passed, requires incapacity be '*because* of an impairment of, or a disturbance in the functioning of, the mind or brain'.[63] These provisions can accommodate some, but not all, cases of undue influence. The need for a 'clear causative nexus between mental impairment and any lack of capacity'[64] leads to a delicate situation when someone's ability to make a decision might be restricted by both an impairment and the influence of another party. If someone with a relevant impairment would be able to decide but for the influence of the other person, then they must be presumed to have mental capacity until all 'practicable steps' are taken without success to counteract that influence.[65] This leaves at least two classes of person who have mental capacity, but might nevertheless be plausibly be thought to lack autonomy. The first are those who 'cannot decide' by the criteria in the Act, but because of undue influence, not because of any impairment. The

[60] *Re T* (n 54) [37].
[61] Law Commission (n 29) para 3.17. This, incidentally, narrows the account of what counts as involuntariness. *Re T* allows that undue influence can also occur when the person can 'think and decide' but is merely attempting 'to satisfy someone else' as a result of their influence (n 53) [31].
[62] Law Commission (n 29) s2(1)(a) (emphasis added).
[63] MCA s2(1) (emphasis added). In the Act, this exclusion of undue influence from the concept of incapacity may reflect Parliament's intentions: Joint Committee on the Draft Mental Incapacity Bill, *First Report* (2002-03, HL 189-II, HC 1083-II) para 270.
[64] *PC* (n 14) [52].
[65] MCA s1(3).

second are those who can decide by the criteria in the Act, but nevertheless cannot reach free decisions because of undue influence. Judges sometimes face both of these sets of circumstances.[66] Consequently, despite the Law Commission's combination of capacity and voluntariness, the insufficiency account has survived, and now coexists with the gatekeeper account. It is found in two places: cases dealing with the inherent jurisdiction, and in the doctrine of undue influence.

3.2 The inherent jurisdiction

The inherent jurisdiction is described by the courts as a 'protective jurisdiction' relating to vulnerable adults,[67] which is available in cases 'that fall outside the MCA'.[68] Not all of its uses are relevant here. For instance, when orders under the jurisdiction are made for people who are incapacitated under the MCA,[69] then this is compatible with the gatekeeper account. Sometimes, however, the jurisdiction is used for people who have mental capacity, and then it relies on an insufficiency account of the relationship between autonomy and capacity.

The leading case, *DL v A Local Authority*,[70] concerned an 'elderly' couple who were thought to be subjected to assaults, coercion, and financial exploitation by their son.[71] The husband lost capacity before the case was heard, but his wife retained capacity throughout.[72] Their local authority sought a wide range of injunctions restraining the son,[73] and the Court of Appeal approved a first instance decision that the inherent jurisdiction allowed orders to be made to protect a 'vulnerable person' who retained capacity, despite the MCA coming into force.[74] Two details must be treated with care here. First, notwithstanding judicial uses

[66] The following discussion largely focuses on cases in which the person could 'decide' under MCA s3(1). When the person cannot 'decide', then the question of causation can be finely balanced. See *NCC v PB* [2014] EWCOP 14, [2015] COPLR 118 [100]-[108]. A notable case decided before the Court of Appeal affirmed the importance of causation in *PC* (n 14) probably treated the distinction between influence and impairment with less care than is required: *Re A (Capacity: Refusal of Contraception)* [2010] EWHC 1549 (Fam), [2011] Fam 61 [65]-[73].
[67] *A Local Authority v SA* [2005] EWHC 2942 (Fam), [2006] 1 FLR 867 [37], [45] (Munby J).
[68] *DL v A Local Authority* [2012] EWCA Civ 253, [2013] Fam 1 [63] (McFarlane LJ).
[69] For matters beyond the jurisdiction granted by the Act, for instance non-recognition of a foreign marriage: *XCC v AA* [2012] EWHC 2183 (COP), [2013] 2 All ER 988. See also *Westminster City Council v C* [2008] EWCA Civ 198, [2009] Fam 11; *A NHS Trust v Dr A* [2013] EWHC 2442 (COP), [2014] Fam 161.
[70] (n 68).
[71] *Ibid* [9].
[72] *Ibid.*
[73] *A Local Authority v DL* [2011] EWHC 1022 (Fam), (2011) 14 CCL Rep 441 [2], [6].
[74] *Ibid* [53]; *DL* (n 68) [68-69].

of the gatekeeper account, the MCA only defines capacity for its own purposes, so it does not exclude people being found 'unable to decide' for reasons other than those found in the Act.[75] This is the legal space that allows the insufficiency account to coexist with the gatekeeper account. Nothing in the MCA requires those with mental capacity to be treated as autonomous. Second, it might be thought that because the injunctions made in *DL* related to the abusive son, they do not bear on whether his mother was recognised to be autonomous. This is not so.

Injunctions like those made in *DL* would normally be made as part of a non-molestation order under section 42 of the Family Law Act 1996. That section, however, does not permit public authorities to initiate proceedings, and the section that does enable third party interventions has never been brought into force.[76] As Miles says, the current situation is usually that 'victims control whether and when legal proceedings for injunctive relief should be brought'.[77] In other words, people who require non-molestation orders are, almost by definition, subject to coercion; but if they have capacity, then they are usually considered to be so autonomous that other parties cannot initiate proceedings for them. Mrs L, and Mr L when proceedings were brought, had capacity. Neither of them initiated proceedings. The use of the inherent jurisdiction changed the situation from the usual one in which they were assumed to be autonomous into one in which their autonomy was in question.[78] This is clear from the judgment itself. Lord Justice McFarlane says that the jurisdiction is 'targeted solely at those adults whose ability to make decisions for themselves has been compromised'.[79] This implies that 'those adults' are not autonomous, a point made explicit when he goes on to say that the jurisdiction is 'aimed at enhancing or liberating the autonomy of a vulnerable adult *whose autonomy has been compromised by a reason other than mental incapacity*'.[80] The possible reasons that a person's autonomy can be 'compromised' are further specified as being 'under constraint', 'subject to coercion or undue influence', or 'for some other reason deprived of the capacity to make the relevant decision or disabled

[75] MCA s2(1); *DL* (n 68) [58]-[61].
[76] Family Law Act 1996 s60.
[77] Jo Miles, 'Family Abuse, Privacy and State Intervention' (2011) 70 Cambridge Law Journal 31, 33.
[78] Whether or not the court truly had the power to do this is questionable: Barbara Hewson, '"Neither Midwives nor Rainmakers" - Why DL Is Wrong' (2013) 3 Public Law 451. *DL* has, however, not been appealed, so it must be treated as reflecting the current state of the law.
[79] *DL* (n 68) [53].
[80] *Ibid* (emphasis added).

from making a free choice, or incapacitated or disabled from giving or expressing a real and genuine consent'.[81] Allowing for the vagueness of the final reason, these are examples of autonomy requiring voluntariness as well as mental capacity.

3.3 The doctrine of undue influence

The inherent jurisdiction has largely been exercised in decisions about a person's health and welfare, but it has a counterpart in decisions about property and affairs: the equitable doctrine of undue influence. In certain circumstances, this doctrine allows the transactions of a person with capacity to be rendered voidable.[82] The leading case, *Royal Bank of Scotland v Etridge (No 2)*,[83] uses slightly different language to *DL*, but its reasoning is analogous. The case consisted of eight conjoined appeals, all by women alleging that their husbands had exercised undue influence when obtaining their agreement to secure loans against matrimonial property. In none of the cases was the mental capacity of the woman in doubt.

Despite the women having capacity, Lord Nicholls held that where undue influence was shown, the transaction 'ought not fairly to be treated as the expression of a person's free will'.[84] Later cases restate his point. For instance, in *Drew v Daniel*, Lord Justice Ward affirms that 'there is no undue influence unless the donor if she were free and informed could say "This is not my wish but I must do it"';[85] and, in *Thompson v Foy*, Mr Justice Lewison held that 'the critical question is whether or not the influence has invaded the free volition of the donor to withstand the influence'.[86] In these cases, as in *DL*, someone with capacity is nevertheless found not to have been in control of their own relevant actions. This falls beyond the gatekeeper account. It is, however, consistent with the insufficiency account, which requires both capacity and voluntariness for autonomy. To be autonomous, both the inherent jurisdiction and the doctrine of undue influence require that someone be free from

[81] *Ibid* [22], [54] approving *SA* (n 67) [77]-[79].
[82] The common law tests for capacity to make a will, an inter-vivos gift, and enter a contract have survived the passing of the MCA. See the Code of Practice (n 33) para 4.31-4.33; *Re Walker* [2014] EWHC 71 (Ch), [2015] COPLR 348 [26]-[50]; *Kicks v Leigh* [2014] EWHC 3926 (Ch), [2015] 4 All ER 329 [37]-[66]. These slightly different tests for capacity do not bear directly on the focus of this chapter: what *follows* from a finding of capacity or incapacity?
[83] [2001] UKHL 44, [2002] 2 AC 773.
[84] *Ibid* [7]. This strand of law usually uses the words 'free will' or 'volition', not 'autonomy', but it is unlikely that anything turns on this difference, given the court's resistance to attempts to be too 'precise or definitive' ([7], [92]).
[85] [2005] EWCA Civ 507, [2005] 2 P&CR DG14 [36].
[86] [2009] EWHC 1076 (Ch), [2010] 1 P&CR 16 [101].

'coercion' and subtler forms of pressure by a third party.[87] Furthermore, in both lines of cases, a special regard must be had for those persons thought to be 'vulnerable', but any attempt to define 'vulnerability' is firmly resisted.[88] In other words, the two strands of law are motivated by the same basic understanding of the relationship between autonomy and mental capacity: one that, as the next subsection discusses, also occurs in some philosophical accounts. On this understanding, autonomy requires both capacity and voluntariness.[89]

3.4 Function of the insufficiency account

After examining the cases in which the insufficiency account appears, two hypotheses can be discarded. First, this is not a case of judicial disagreement. It is not that some judges rely on one account and other judges rely on another account. Lord Justice McFarlane gave the leading judgment in both *DL* and *PC*, but the former is premised on an insufficiency account and the latter on the gatekeeper account. Similarly, Lady Justice Butler-Sloss gave judgment in both *Re T*, which relied on the insufficiency account, and in *Re MB*, which relied on the gatekeeper account. Second, no simple story about the evolution of judicial views will work here. An insufficiency account in *Re T* predates a gatekeeper account in *MB*, which predates an insufficiency account in *DL*, which predates a gatekeeper account in *PC*. Rather than differences between judges or over time, the differences between the accounts are better accounted for by paying attention to the differences between the cases in which they are used. The previous section discusses how the gatekeeper account can add rhetorical force to a judgment when someone with capacity is making an unwise choice or when the best interests of someone without capacity are felt to contradict their expressed wishes. If, however, a third party is coercing someone with capacity, then a gatekeeper account would often seem to exclude any action by the court. After all, if someone with capacity is always so self-governing that the state should not interfere with their expressed wishes, then even wishes expressed at gunpoint should be respected as autonomous. This is not a particularly attractive conclusion, nor is it one that is likely to persuade; so it is hardly surprising that when coercion may be an issue, judges instead give an insufficiency account. In doing so,

[87] *DL* (n 68) [61]; *Etridge* (n 83) [8].
[88] *DL* (n 68) [63]-[64]; *Etridge* (n 83) [11], [18].
[89] This does not imply that all details of the interaction between the inherent jurisdiction and the equitable doctrine are legally settled. See *OH v Craven* [2016] EWHC 3146 (QB) [36].

they draw a distinction between the internal and external limitations on autonomy, one that also appears in the philosophical literature. For instance, Raz distinguishes between 'capacity' and 'independence',[90] and Killmister between 'capacity for autonomy' and 'ability for autonomy'.[91] For Killmister, the former assesses someone's potential for autonomy in ideal circumstances, and the latter their potential for autonomy in their current circumstances. The gatekeeper account emphasises 'capacity', in Killmister's sense, at the expense of 'ability'. The MCA requires an inability to decide be because of an impairment,[92] so it is largely restricted to the internal constraints on autonomy; and the gatekeeper account, by tying autonomy to capacity in the Act, adopts these limitations.[93] In contrast, the insufficiency account, which says that autonomy requires both capacity and voluntariness, allows attention to the person's external circumstances. From this description, it might be thought that the gatekeeper account is just a slightly inaccurate shorthand for the insufficiency account; and that when coercion does not seem to be a factor, judges just do not mention it. There is some truth to this, but it would be wrong to conclude that the insufficiency account accurately reflects the whole law. In contrast to what it would predict, judges sometimes hold that a person without capacity is autonomous.

4. The survival account

On the survival account, unlike the gatekeeper and insufficiency accounts, respect for the autonomy of a person without mental capacity can restrain state action. Like the other accounts, it has been bluntly stated by the judiciary. For instance, in *W v M,* Mr Justice Baker held that 'personal autonomy survives the onset of incapacity'.[94] This statement is unequivocal, but the position it articulates is less simple than it appears. It should be distinguished from two opposing extremes. At one extreme is the view that all people, regardless of mental capacity, are so autonomous that the state should not act against their will.[95] At the other extreme is the view that someone without mental capacity might have

[90] Joseph Raz, *The Morality of Freedom* (Oxford University Press 1986) 378-390.
[91] Suzy Killmister, 'Autonomy, Liberalism, and Anti-Perfectionism' (2013) 19 Res Publica 353, 358. Similarly, Maclean distinguishes 'voluntariness' and 'competence': Alasdair Maclean, *Autonomy, Informed Consent and Medical Law: A Relational Challenge* (Cambridge University Press 2009) 139-40.
[92] MCA s2(1); *PC* (n 14) [59]-[61].
[93] This point is sometimes made by saying that the MCA favours a 'medical model' of incapacity to a 'social model'. Beverly Clough, 'People Like That: Realising the Social Model in Mental Capacity Jurisprudence' (2014) 23 Medical Law Review 53, 56-59.
[94] [2011] EWCOP 2443, [2012] 1 WLR 1653 [95].
[95] This 'universal account' is discussed in subsection 5.2.

some capability for self-determination, but does not have a legal right to be treated as autonomous. This position, as noted in section two, collapses into a gatekeeper or insufficiency account. In contrast, the survival account asserts that a right to be treated as autonomous *can* survive the onset of incapacity, but not that it *must* do so. This is clear when Mr Justice Baker goes on to say:

> A decision by the Court, having proper regard to the patient's personal autonomy and the expressed wishes and feelings of the patient and her family, that it would be in her best interests to withhold or withdraw treatment does not give rise to a breach of Article 8.[96]

This suggests that appeals to respect for autonomy, arguments that a person is so self-governing that their expressed wishes must be followed, might be effective even after someone is found to lack capacity. In other words, even when someone is found to lack capacity, respect for their autonomy can restrict the decisions that can be made on their behalf. The question of respect for autonomy is not resolved by a capacity assessment, it is determined when evaluating a person's best interests.

4.1 Recent examples of the survival account

Section two showed how the gatekeeper account draws some plausibility from the Mental Capacity Act 2005. The Act, however, clearly envisages that a person without capacity might be able to at least contribute to decisions made on their behalf. When a best interests decision is being made, the person being decided for should be permitted and encouraged to participate in decision making;[97] and their 'past and *present* wishes and feelings' must be considered.[98] Similarly, the Act requires that 'regard must be had' to making decisions in a

[96] *W v M* (n 94) [95]. 'Article 8' refers to that part of the European Convention on Human Rights, which protects personal autonomy: *Pretty v UK* (2002) 35 EHRR 1 [61]; *Evans v UK* (2008) 46 EHRR 34 (GC) [71]. Mr Justice Baker arguably oversimplifies the law here. The Court of Appeal has since held that the 'safe approach' is for the court to do a full best interests assessment under the MCA *and then* to also assess whether any apparent interference with the person's Article 8 rights is 'necessary and proportionate': *K v LBX* [2012] EWCA Civ 79, [2012] 1 FCR 441 [35] (Thorpe LJ). This two-step process is not consistently followed in the Court of Protection, even when citing *LBX*: for instance, *ACCG v MN* [2013] EWHC 3859 (COP), [2014] COPLR 11 [81]-[85]. A fuller treatment of this issue is beyond the remit of this chapter; for, whatever the differences between treating human rights as part of best interests or as an additional test, both approaches are premised on a survival account of the relationship between autonomy and mental capacity.

[97] MCA s4(4).

[98] MCA s4(6)(a) (emphasis added). Section 4(6)(b), concerning 'beliefs and values', addresses those that the person would have if they had capacity. It does not directly bear on whether a person without capacity can retain such autonomy that the court should treat their expressed decision as final. Sometimes decisions are

way that is 'less restrictive of the person's rights and freedom of action'.[99] These sections require more attention to someone's preferences than was evident in the pre-Act case law.[100] They do not, however, straightforwardly translate into any right to be treated as autonomous. Someone's wishes and feelings must only be considered, not followed, and they have no priority over other factors.[101] Similarly, the Act only requires regard to what is 'less restrictive'. It does not require decision-makers to act in the least restrictive way unless that is in the person's overall best interests.[102] In theory, then, the Act requires consideration of the person's point of view, but does not require treating the person without capacity as so autonomous that the state should not act against their will. In practice, however, attention to the person's point of view has tended to dissolve this distinction.

There are two leading cases on the current 'wishes and feelings' of a person without capacity. The first, *Re P (Statutory Will)*, establishes that, although they carry 'great weight', 'the only imperative is that the decision must be made in P's best interests'.[103] This adds little to the plain words of the Act, but the second case, *ITW v Z*,[104] fills in more detail. In it, Mr Justice Munby noted that the weight attached to someone's wishes and feelings is always case-specific,[105] then gave a non-exhaustive list of factors bearing on the issue:

> a) the degree of P's incapacity, for the nearer to the borderline the more weight must in principle be attached to P's wishes and feelings…
>
> b) the strength and consistency of the views being expressed by P;

---

made based on a person's past wishes or values, and this is thought to protect autonomy. If, however, the person in question is expressing a *present* preference to the contrary, then this is deeply suspicious. Burch is right that such decisions 'thereby deny her the right to make a decision on the grounds that they understand who she is better than she does. The claim that this is not a coercive denial of P's legal capacity is a kind of doublespeak where coercion doesn't always mean coercion': Matthew Burch, 'Autonomy, Respect, and the Rights of Persons with Disabilities in Crisis' [2016] 34 Journal of Applied Philosophy 389.

[99] MCA s1(6).

[100] Mary Donnelly, 'Best Interests, Patient Participation and the Mental Capacity Act 2005' (2009) 17 Medical Law Review 1.

[101] The Law Commission has suggested a stronger duty to ascertain a person's wishes and that they be given 'particular weight' among the factors evaluated in a best interests decision. Law Commission, *Mental Capacity and Deprivation of Liberty* (Law Com No 372, 2017) para 14.10-14.21, Recommendation 40. This would not change the nature of the survival account as it appears in this section. It may, however, require decision-makers to rely on this account more frequently.

[102] *C v A Local Authority* [2011] EWHC 1539 (Admin), [2011] Med LR 415 [61](Ryder J).

[103] [2009] EWHC 163 (Ch), [2010] Ch 33 [41] (Lewison J). See also *Re CD* [2015] EWCOP 74, [2016] COPLR 1 [28] (Mostyn J).

[104] [2009] EWHC 2525 (Fam), [2011] 1 WLR 344.

[105] *Ibid* [35].

> c) the possible impact on P of knowledge that her wishes and feelings are not being given effect to…
>
> d) the extent to which P's wishes and feelings are, or are not, rational, sensible, responsible and pragmatically capable of sensible implementation in the particular circumstances; and
>
> e) crucially, the extent to which P's wishes and feelings, if given effect to, can properly be accommodated within the court's overall assessment of what is in her best interests.[106]

The first and last of these factors directly bear on the structure of the relationship between autonomy and capacity. The remaining three do so indirectly, by directing attention to the content of a person's wishes. The first factor, stating that more weight should be given to the wishes and feelings of people closer to the borderline of capacity, may seem obvious, but even it tends to undermine the gatekeeper account. Instead of a simple binary where those with capacity have autonomy and those without capacity do not, autonomy becomes a matter of degree; and if you are close to the 'borderline', then you are also closer to your expressed wishes being legally determinative.[107]

The other structural factor from *ITW*, the requirement that a person's wishes can 'be accommodated within the court's overall assessment of what is in her best interests', may seem to limit how far the gatekeeper account can be softened. If someone's expressed feelings are not in accord with their overall best interests, then their wishes will not prevail; so it can seem that, after all, only those with capacity have a right to be treated as autonomous. This appearance is deceptive, and anyone led astray by it has failed to ask whether the wishes and feelings of a person without capacity are a factor in determining their overall best interests. They are. As already mentioned, the Act requires that the person's present wishes are considered;[108] and sometimes this consideration leads to a judge treating those wishes as determinative of the person's objective best interests.[109] For

---

[106] *Ibid*.

[107] Herring argued for this development in an early article on the Act that read it through the prism of the gatekeeper account. His article predates almost all of the case law discussed in this section. Jonathan Herring, 'Losing it? Losing what? The Law and Dementia' (2009) 21 Child and Family Law Quarterly 3.

[108] MCA s4(6).

[109] Judge Hazel Marshall QC rhetorically asks 'What, after all, is the point of taking great trouble to ascertain or deduce P's views, and to encourage P to be involved in the decision-making process, unless the objective

example, in *Re EU*, Senior Judge Lush found that 'The factor of magnetic importance in this case is EU's own wishes and preference'.[110] In *Wye Valley v B*, Mr Justice Peter Jackson held that 'the wishes and feelings, beliefs and values of a person with a mental illness can be of such long standing that they are an inextricable part of the person that he is',[111] and this determined best interests in the case.[112] Finally, in *SAD v SED*, District Judge Glentworth found 'the Respondent's [current] wishes and feelings are central to my decision'[113] when abiding by the desire of someone without capacity to revoke a Lasting Power of Attorney that they had made when they had capacity. Overall best interests cannot entirely limit the determinative power of someone's current wishes because sometimes the person's current wishes determine their overall best interests.[114] This may seem a modest point, but it cannot be accounted for by the gatekeeper and insufficiency accounts, both of which require mental capacity for a person's autonomy to restrain interference with their expressed wishes.

4.2 Subjectivity, intelligibility, and authenticity

In *ITW,* beyond the two points that directly address the relationship between capacity and autonomy, are three factors that indirectly address the relationship. Each of these expresses a conception of autonomy that is not easily assimilated into the gatekeeper view. The three conceptions are subjectivity, intelligibility, and authenticity; and the first of these, subjectivity, is expressed as 'the possible impact on P of knowledge that her wishes and feelings are not being given effect to'. The issue here is simple. Best interests decisions are made for people, not objects, and people have their own perspectives on the world. As District Judge Eldergill says, the 'law requires objective analysis of a subject not an object'.[115] The Supreme Court agrees. In *Aintree v James,* Lady Hale held that the 'purpose of the best interests test is to consider matters from the patient's point of view'.[116] In *Aintree* itself, Mr

is to try to achieve the outcome which P wants or prefers, even if he does not have the capacity to achieve it for himself?' *Re S (Protected Persons)* [2010] 1 WLR 1082 (COP) [55]. Cited with approval for 'all applications of the best interests test' by Charles J in *Briggs v Briggs* [2016] EWCOP 53, [2017] 4 WLR 37 [59].

[110] [2014] EWCOP 21 [46]. Similarly, see *Re GW* [2014] EWCOP B23 [41].

[111] *Wye Valley NHS Trust v B* [2015] EWCOP 60, [2015] COPLR 843 [13].

[112] *Ibid* [42].

[113] [2017] EWCOP 3 [29].

[114] It is an open question how far this point is appreciated in practice outside the Court of Protection, but with regards past wishes some doctors are aware of this possibility: *Abertawe Bro Morgannwg University LHB v RY* [2017] EWCOP 2, [2017] COPLR 143 [40].

[115] *Westminster City Council v Sykes* [2014] EWHC B9 (COP), (2014) 17 CCL Rep 139 §10.

[116] (n 4) [45].

James's current wishes did not play a prominent role,[117] but attention to the person's 'point of view' has naturally led to their wishes being given weight in other cases. For instance, in *A NHS Foundation Trust v Ms X*, Mr Justice Cobb 'endeavoured to put myself in the place of Ms X'.[118] In doing so, he carefully considered her current wishes and the likely effects of overruling them.[119] Similarly, in *Wye Valley,* Mr Justice Peter Jackson observed that if he found it in a man's best interests to have a severely infected leg amputated against his will, then 'the loss of his foot will be a continual reminder that his wishes were not respected',[120] and this counted against the operation being performed. Sometimes, for instance in *Ms X*,[121] these concerns are linked to autonomy, but the nature of this connection has never been made entirely clear. In an older case concerning someone found to have capacity, however, the President of the Family Division did give some indication of the connection between subjectivity and autonomy.[122] She quoted at length an article by the philosopher Kim Atkins, including the following passage:

> If we accept that the subjective character of experience is irreducible and that it is grounded in the particularity of our points of view, then we are bound to realise that our respect for each other's differences and autonomy embodies a respect for the particularity of each other's points of view.[123]

It might be thought that the word 'respect' in this sentence is relatively weak, and that 'respect for the particularity' of someone's point of view does not entail that their wishes should necessarily prevail. This conclusion should be resisted. In *Ms B,* this respect for autonomy entailed allowing her to make a decision that would almost certainly result in her own death.[124] It seems likely that cases such as *Ms X* and *Wye Valley* now apply a parallel

[117] At first instance, however, his likely wishes, insofar as they could be ascertained, were taken into account: *An NHS Trust v DJ* [2012] EWCOP 3524 [79].
[118] [2014] EWCOP 35, [2015] COPLR 11 [58].
[119] *Ibid* [44]-[54].
[120] (n 111) [37]. Similar reasoning was critical in a decision to allow a man with an acquired brain injury travel to Serbia for an experimental stem cell treatment with very little evidence of its effectiveness: *Re D (Medical Treatment)* [2017] EWCOP 15, [2017] COPLR 347 [60].
[121] (n 118) [2], [44], [46].
[122] *Ms B v An NHS Hospital Trust* [2002] EWHC 429 (Fam), [2002] 2 All ER 449 [81]-[83] (Butler-Sloss P)
[123] Kim Atkins, 'Autonomy and the Subjective Character of Experience' (2000) 17 Journal of Applied Philosophy 71.
[124] (n 122) [94]-[95].

line of thought to some people without capacity;[125] and in these cases, too, respecting the wishes of the person risked their death. Attention to subjectivity increases the likelihood of a person's wishes determining their objective best interests, for considering things from the person's point of view tends to make their wishes central, instead of merely one view among many. When the person's perspective determines their best interest in this way, a right to autonomy has survived incapacity, and restrained third parties from acting on their behalf.

Intelligibility, 'the extent to which P's wishes and feelings are, or are not, rational, sensible, responsible and pragmatically capable of sensible implementation in the particular circumstances', is the second factor from ITW that addresses the content of the person's wishes. Once again, this factor tends to undermine the gatekeeper account. If someone without capacity can be expressing a 'rational, sensible, responsible' preference, this calls into question the idea that those without capacity lack autonomy.[126] The Court of Protection can be open to people without capacity having good reasons for their preferences. For instance, in *A London Borough v VT,* District Judge Ralton found that ST's wish to return to Nigeria 'to die in the land of his birth'[127] was 'an expressed consistent rational intention',[128] one that he shared with many peers with capacity.[129] This is an important detail. People do not make sense of one another in a void. In everyday life, they interpret one another by referring to their shared social context, social practices, and personal narratives;[130] and when interpreting the preferences of someone without capacity, they necessarily start in the same way. *Ms X* provides another example. In this case, Mr Justice Cobb considered the wishes of a woman with anorexia nervosa who refused life-sustaining nutrition despite having 'no wish to die'.[131] At first, such a decision might seem incomprehensible, but the

[125] A detailed examination of the role of the 'subjective character of experience' in best interests decisions can be found in Laura Pritchard-Jones, '"This Man with Dementia"—"Othering" the Person with Dementia in the Court of Protection' (2016) 24 Medical Law Review 518.
[126] This development in the treatment of incapacity parallels a broader change in attitudes to mental health. As Glover says, it used to be thought that the inner lives of people with mental disorders were 'impenetrably alien' and beyond interpretation by others; but this idea is increasingly under challenge as efforts are made to find the 'right interpretation'. Jonathan Glover, *Alien Landscapes? Interpreting Disordered Minds* (Harvard University Press 2014) 1, 126-127.
[127] [2011] EWCOP 3806 [35].
[128] *Ibid* [37], [41].
[129] *Ibid* [35].
[130] Alasdair MacIntyre, 'The Intelligibility of Action' in J Margolis (ed), *Rationality, Relativism and the Human Sciences* (Martinus Nijhoff 1986) 63.
[131] (n 118) [38].

judge rendered it intelligible by describing Ms X's long history of unsuccessful similar admissions to specialist units.[132] By describing what her experiences were like, he made her apparently contradictory wishes part of a coherent, understandable, story. This attention to someone's 'narrative' is not unique. It appears in an increasing number of cases,[133] and telling someone's story in this way often renders their wishes intelligible. Intelligible wishes are far more likely than unintelligible ones to determine overall best wishes; so, as with subjectivity, this process tends to add weight to the survival account. When, as in *VT*, someone without capacity is described as having 'an expressed consistent rational intention' and that intention decisively determines their best interests,[134] then the language of autonomy has invaded the land of incapacity.

The final factor from *ITW*, 'the strength and consistency of the views being expressed by P', is authenticity. This reflects an idea often associated with autonomy. For philosophers, autonomy is often taken to require a second-order identification with one's own desires, or at least that the person would not feel alienated from their own preferences if they happened to critically reflect on them.[135] In many ways these theories, which point towards internal coherence, merely formalise an everyday practice. When, in the normal course of things, people interpret one another's behaviour, they sometimes rely on the idea that the more strongly or consistently someone expresses a desire, the more that they 'really' want that thing. Things that are authentically wanted in this way are usually considered more important to the person than things that are not. Similar reasoning occurs in the cases.[136] *Wye Valley* is particularly striking:

> …the wishes and feelings, beliefs and values of a person with a mental illness can be of such long standing that they are an inextricable part of the person that he is. In this situation, I do not find it helpful to see the person as if he were a

[132] *Ibid* [12]-[25].
[133] Carolyn Johnston and others, 'Patient Narrative: An "On-Switch" for Evaluating Best Interests' (2016) 38 Journal of Social Welfare and Family Law 249.
[134] (n 127) [41], [68].
[135] John Christman, *The Politics of Persons: Individual Autonomy and Socio-Historical Selves* (Cambridge University Press 2009) ch 7.
[136] Beyond the existing practice described here, the minor reforms to s4(6) of the Act suggested by the Law Commission are intended to especially protect strong and clear wishes: Law Commission (n 101) para 14.17. More extreme is the suggestion of a legal standard based entirely on authenticity instead of best interests mentioned in Wayne Martin and others, *Three Jurisdictions Report: Towards Compliance with CRPD Art 12 in Capacity/Incapacity Legislation across the UK* (University of Essex 2016) 41-42.

> person in good health who has been afflicted by illness. It is more real and more respectful to recognise him for who he is: a person with his own intrinsic beliefs and values.[137]

In this case, even acknowledging that Mr B's wishes followed from 'delusions arising from mental illness' did not mean that they should be given any less weight.[138] Despite their apparent irrationality, Mr B's desires were an 'inextricable' and 'intrinsic' part of him, and so worthy of respect. Similarly, in *SAD v SED*, District Judge Glentworth found of a woman with bipolar affective disorder, 'it cannot be said that her wishes and feelings can be delineated in a way which gives more weight to those expressed in a phase when the hypomania is absent than when …[it] affects her behaviour'.[139] Both her symptomatic and asymptomatic selves were 'part of who she is'.[140] In these cases, the authenticity of someone's wishes and feelings allows them to determine that person's overall best interests. As with subjectivity and intelligibility, authenticity allows autonomy to survive incapacity.

4.3 Function of the survival account

Sometimes, after considering a person's subjectivity and the intelligibility and authenticity of their wishes and feelings, a judge in the Court of Protection will find that those wishes and feelings determine the person's objective best interests. This is not consistent with the gatekeeper and insufficiency accounts. In these cases, someone without capacity demonstrates that they are so capable of self-determination that the state should not interfere with their decision. An appeal to respect for autonomy is successful. This conclusion might provoke some resistance. Restricting a judge to abiding by someone's wishes only when it is in that person's overall best interests might not be thought strong enough to be called the law 'respecting autonomy'. Such resistance would be misplaced. Having a legal right that is subject to conditions (here, the person's overall best interests) does not mean that there is no legal right. It means that there is a conditional legal right.[141] In capacity law, this is not unique. If people with capacity have a legal right to make autonomous decisions, then that right is dependent on them being presumed or found to

[137] (n 111) [13] (Peter Jackson J).
[138] *Ibid*.
[139] (n 113) [28].
[140] *Ibid*.
[141] Andrew Melnyk, 'Is there a Formal Argument against Positive Rights?' (1989) 55 Philosophical Studies 205; Michael Levin, 'Conditional Rights' (1989) 55 Philosophical Studies 211.

have capacity. It, too, is a conditional right. It is true that best interests decisions leave a lot of discretion to the decision-maker, but so do capacity assessments, and in the Court of Protection they are performed by a small group of expert witnesses.[142] It is also true that there are reasons to doubt whether best interests decisions are consistently made with due rigour,[143] but there are corresponding reasons to doubt the consistency of capacity assessments.[144] Whatever the state of practice, in law autonomy can survive incapacity.[145]

It is easy to use the survival account to tell a comforting story. In this story, the law is progressively realising the rights of people without capacity by recognising autonomous decisions where before it only saw irrational desire. This kind of fairy tale is better avoided. Since the passing of the MCA, the law now pays more attention to the wishes of the person who is being decided for. This is admirable, but it is not necessarily empowerment. Control remains with whoever is taken to be the authoritative interpreter of the person's wishes. This is not the person themselves. It is a wide range of professionals and, ultimately, a judge.[146] Furthermore, attention to someone's wishes requires finding out what those wishes are. For some people without capacity, this might require intrusive questioning;[147] and some may feel that strangers, no longer satisfied with merely monitoring their outward behaviour, are now prying into their mental life. Although it is possible to tell a story about how attention to wishes and feelings will empower people, it is also possible to tell a story in which it leads to a new form of domination. In the darker story, authenticity slowly becomes the central criteria for decision-making, and professionals are freed from just

[142] Paula Case, 'Negotiating the Domain of Mental Capacity: Clinical Judgement or Judicial Diagnosis?' (2016) 16 Medical Law International 174.

[143] A particular worry is whether the wishes and feelings of people with learning disabilities are given as much weight as other groups: Lucy Series, 'The Place of Wishes and Feelings in Best Interests Decisions' (2016) 79 Modern Law Review 1101. This is consistent with, and even suggested by, the account given here. Unlike, for instance, someone with a progressive condition, such as dementia, a person with a learning disability may not have an earlier period of capacity that autonomy can be characterised as having *survived from*.

[144] Select Committee on the Mental Capacity Act 2005, *Mental Capacity Act 2005: Post-Legislative Scrutiny* (2013-14, HL Paper 139) para 60.

[145] This is not intended to minimise the problems implementing the Act, which are extensive. *Ibid.*

[146] Those who advocate moving even further towards approaches based on authenticity sometimes expressly endorse the point that others retain interpretative control. For instance: G Szmukler and M Bach, 'Mental Health Disabilities and Human Rights Protections' (2015) 2 Global Mental Health e20, 7. Although the final decision is judicial, Clough points out that even cases where the judge pays close attention to the person's wishes are often deeply influenced by medical opinion: Beverly Clough, 'Anorexia, Capacity, and Best Interests: Developments in the Court Of Protection since The Mental Capacity Act 2005' (2016) 24 Medical Law Review 434, 444.

[147] This is already an issue. For instance, *Re SL* [2017] EWCOP 5 [12] (Glentworth DJ): 'She is described as regretting that she has to contribute to the cost, that it has gone on so long and she does not like having to talk to people about the matters repeatedly'.

dictating what individuals in a 'benighted state' can do.[148] Instead, they also dictate what those people 'really' want, whether or not the people in question know it.[149] Even in this darker future, the process will undoubtedly still be called 'empowerment'; but, in truth, it will offer nothing but the freedom of a political re-education camp.[150]

Rather than reading cases through a narrative of future empowerment, it is better to focus on their current effect. Doing so highlights an important detail. The survival account has not replaced the other accounts. It coexists with them, but is used in different circumstances. When the facts mean that it would be artificial to hold that someone can understand or 'use or weigh' the information relevant to a decision, yet they seem relevantly autonomous, it allows the judge to recognise this broader type of self-determination. *Wye Valley* is prototypical. It would have been difficult for Mr Justice Peter Jackson to hold that a man with treatment-resistant schizoaffective disorder who did not 'understand the reality of his injury', and became agitated when it was even discussed, had capacity to make decisions about treatment for that injury.[151] The survival account nevertheless allowed him to abide by the man's intelligible and authentically held wishes.[152] To hold that autonomy survives incapacity seems to contradict the gatekeeper and insufficiency accounts, just as they seem to contradict one another. This does not, however, mean that the relationship between autonomy and mental capacity is incoherent. It is not. Rather, the courts emphasise the particular elements of the relationship that they find relevant in the case before them, and this has led to three partial accounts of a coherent whole.

5. A coherent account

In *W v M,* Mr Justice Baker says that 'personal autonomy survives the onset of incapacity';[153] but in *Bailey v Warren,* Lady Justice Arden says that 'capacity is an important issue because

[148] Berlin (n 11) 205.
[149] Cases that emphasise someone's past wishes or the values that they would be likely to have if they had capacity instead of the person's *present* wishes already show elements of this attitude. For instance, such values and past feelings would have justified withdrawing artificial nutrition and hydration to a man with severe brain injuries even if he had currently demonstrated 'contentment and happiness': *Briggs* (n 109) [44]-[52], [117]-[119], [129]-[130]. For further comment on this tendency, see Burch (n 98).
[150] A similar dynamic already occurs under the MCA, which is made to fit 'clinical judgments', 'resource-led decision-making', or 'the safeguarding agenda' instead of modifying these things: Select Committee (n 144) para 104.
[151] (n 111) [34].
[152] *Ibid* [42]-[45].
[153] (n 94) [95].

it determines whether an individual will in law have autonomy over decision-making'.[154] In *DL*, McFarlane LJ talks about the 'adult whose autonomy has been compromised by a reason other than mental incapacity'; [155] but, in *PC,* he affirms that 'unless they lack mental capacity …the statute respects their autonomy'.[156] It would be easy to read these statements and conclude that there is no coherent relationship between autonomy and mental capacity in domestic law. Some might even conclude that 'autonomy' is simply used to excuse decisions made on other grounds, and that a person's mental capacity or incapacity is only as significant the judge chooses to make it. Such cynicism should be resisted. It does not fit the evidence. All of the cases discussed here give an impression of anxious moral scrutiny, not of sophistry or arbitrary choice. Furthermore, each account has characteristic uses. Each reliably appears in response to certain types of concern; and these concerns, for instance coercion by a third party, are things that judges should be responsive to. More fundamentally, cheap cynicism makes the same sort of oversimplification as the three accounts.

5.1 A coherent account

The cynic and three previous accounts all make the same simple but common error. Midgley describes it as expecting a 'single plain litmus-paper test' for a complex moral concept.[157] The gatekeeper account gives the impression that there is such a simple test for the concept of autonomy: if you have capacity then you are autonomous, and if you do not then you are not. The insufficiency and survival accounts each complicate one limb of the gatekeeper account, but neither abandons the ideal of a single test for autonomy. In the insufficiency account, the single test is whether the person has both capacity and freedom, and in the survival account the test is part of the best interests assessment. The cynic rightly observes that there is no single test, and wrongly infers that there is no moral concept. As with the three accounts, the assumption is that unless the test is relatively simple, there is no such thing as autonomy. Instead, it is better to recognise with Midgley that moral concepts can lack any simple test, yet have an 'underlying structure' in practical reasoning.[158] To tease out the underlying structure of autonomy in mental capacity law, a

[154] (n 40) [105].
[155] (n 68) [54].
[156] (n 14)[53].
[157] Mary Midgley, 'The Game Game' in *Heart and Mind* (first published 1974, Routledge 2003) 154, 184.
[158] *Ibid*.

narrow focus on what a judge says in any one case must be replaced with a descriptive account of how the concept is used across a range of cases. On this broader view, a coherent account of the relationship between mental capacity and autonomy is visible. To the question, 'when is a person so capable of self-government that the state will not interfere with their wishes?' the answer is as follows:

> **In civil law, an adult in England and Wales has a legal right to be treated as autonomous with regard to a particular matter if they are not detained under the Mental Health Act 1983 and:**[159]
>
> **(1)** They have capacity **and** are not a vulnerable person subject to coercion or undue influence,[160] **or**
>
> **(2)** They lack capacity **but** the character of their wishes and feelings is such that it determines their objective best interests.[161]

Mental capacity, then, is not the gatekeeper to autonomy, but it is an important part of the legal threshold for autonomy as a right. Rather than determining the question of whether a person should be treated as autonomous, the concept of mental capacity creates two presumptions. If a person is found to have capacity, then they will be presumed to be autonomous, but that presumption may be rebutted if they are found to be vulnerable and subject to coercion. If a person is found to lack capacity, then they will be presumed not to be autonomous, but that presumption will be rebutted if their wishes and feelings show such subjectivity, intelligibility, and authenticity that they must be respected.[162]

---

[159] The clause 'not detained under the Mental Health Act 1983' indicates that this topic is beyond the scope of this chapter, ***not*** that persons so detained can necessarily be treated as lacking in autonomy. Although the Act allows detention for the patient's 'health and safety' (ss 2, 3) and involuntary treatment 'for the mental disorder' (s63) regardless of capacity, it does not grant any general power to treat a person as lacking autonomy. For instance, the ability to treat does not extend to unconnected physical disorders and has exceptions (ss57, 58, 58A). Furthermore, the Act's Code of Practice makes the 'least restrictive option and maximising independence' and 'empowerment and involvement' overarching principles; so, insofar as the Act generates exceptions to the law outlined here, they should be narrowly constructed. Department of Health, *Mental Health Act 1983: Code of Practice* (The Stationery Office 2015) paras 1.1-1.12.

[160] See section 3. *DL* (n 68) [54], [64]; *Etridge* (n 83) [8]-[18].

[161] See section 4. *ITW* (n 104) [35]. Any finding that a person's wishes and feelings do not have this character may be subject to an additional test of necessity and proportionality under Article 8 of the European Convention on Human Rights (n 96). It seems likely that when someone without capacity is subject to coercion or undue influence, then their wishes and feeling will not be determinative of their objective best interests.

[162] If the recent Law Commission proposals (n 101) were implemented this latter presumption would be effectively reversed to a presumption that those without capacity *are nevertheless autonomous* unless their wishes and feelings *fail to* show such subjectivity, intelligibility, and authenticity that they must be respected.

5.2 Implications for the Mental Capacity Act 2005

The coherent account of the relationship lacks the simplicity of the three accounts that judges have relied on. In many ways, this is down to a difference of purpose. Judges are concerned with those aspects of the relationship between autonomy and capacity that are relevant in the particular case, but this chapter is focussed on the relationship as a whole. This bird's eye view may, however, have implications for judicial practice, without requiring legal reform. The coherent account lends itself naturally to being used as a decision tree. First, determine whether the person has capacity. If they do, consider whether they are subject to undue influence and coercion: if they are not coerced, then they may act 'for reasons which are rational, or irrational, or for no reason';[163] if they are coerced, then measures to address the coercion may be available under the inherent jurisdiction. If the person lacks the relevant capacity, then their wishes must be taken seriously when determining their best interests, and questions of undue influence and coercion should still be considered.

This outline of the relationship between autonomy and mental capacity is both coherent and relatively determinate. It too, however, should avoid the temptation to posit a simple test for a complex concept; for in the details, indeterminacy appears. If someone is truly acting for reasons that are 'irrational', then will they be found to have capacity? In many cases, it is unlikely; so the, often-quoted, passage from *Sidaway* that asserts a right to act this way must be treated as another example of judicial rhetoric,[164] one that may work to great effect in particular cases while distorting a realistic assessment of the law as a whole. Furthermore, this chapter takes both capacity and undue influence as relatively unquestioned matters of fact, but any judge must also decide what counts as 'understanding' the relevant information or what makes influence 'undue' in the particular circumstances of a case. In other words, they do not merely apply a test to the facts. In every case, they are partially responsible for the content of the test itself.

Having judges make the legal standards as they apply them is an old response to an old problem, one noted by Aristotle: 'all law is universal, and yet there are some things about

[163] *Sidaway v Board of Governors of the Bethlem Royal Hospital* [1985] UKHL 1, [1985] AC 871, 904.
[164] *Ibid*.

which it is not possible to make correct universal pronouncements'.[165] Any determinate test risks injustice, but to acknowledge an unending list of exceptions, and exceptions to exceptions, risks losing the rule in an avalanche of particular cases. An appreciation of this tension was incorporated into English law, through the principle of equity, by the Middle Ages; and at some points, for instance during the sixteenth century, it has been the central issue in legal thought.[166] In the modern era, however, its prominence has faded, and judicial discretion often meets academic suspicion.[167] Nevertheless, there is a parallel between the traditional idea that equity mitigates the inflexibility of the common law and the development of capacity law since the passing of the MCA. It would have been open to the court in *DL* to find that the inherent jurisdiction did not survive the passing of the Act, and in *ITW* to hold that the requirement to 'consider' a person's wishes as part of a best interests decision never entailed that those wishes determine the person's objective best interests. If the courts had taken these severe lines, then the gatekeeper account would have been true. They did not, and the inflexibility of that account has been mitigated.

The parallel with the principle of equity cannot be drawn too closely. Many of the post-MCA cases have precedential force, so the coherent account is not merely a description of how discretion is exercised. It is hardening into binding law. This raises a serious question. One rationale for the MCA was that it would simplify an area of law that had become 'confusing, often incomplete, and much misunderstood'.[168] This chapter excludes children, the criminal law, and those subject to the Mental Health Act 1983;[169] yet, even with those restrictions, determining whether an individual should be treated as autonomous requires considering the MCA, the inherent jurisdiction of the High Court, the equitable doctrine of undue influence, and the list of factors found in *ITW.* Furthermore, this chapter discusses only the consequences of a finding of capacity or incapacity. It treats the actual assessment

[165] Aristotle, *Nicomachean Ethics* (first published approximately 322 BC, Christopher Rowe tr, Oxford University Press 2002) 1137b13.

[166] Stuart E Prall, 'The Development of Equity in Tudor England' (1964) 8 The American Journal of Legal History 1.

[167] Writers with judicial experience still tend to see the tension between universal laws and justice in the particular case as important and inescapable. For example, see Thomas (n 2) 116-117; Hedley (n 3) 58-60.

[168] HL Deb 10 January 2005, vol 668, col 13 (Lord Falconer LC).

[169] Cave discusses the inconsistencies and gaps between the MCA, the MHA, and laws pertaining to children. Emma Cave, 'Determining Capacity to Make Medical Treatment Decisions: Problems Implementing the Mental Capacity Act 2005' (2014) 36 Statute Law Review 86. Peay gives an account of the challenges facing the relevant areas of criminal law: unfitness to plead, automatism, and the verdict of not guilty by reason of insanity. Jill Peay, 'Mental Incapacity and Criminal Liability: Redrawing the Fault Lines?' (2015) 40 International Journal of Law and Psychiatry 25.

process as unproblematic. In reality, capacity assessments face a variety of legal uncertainties, tangles, and paradoxes.[170] Most people applying the Act are not legally trained, so it seems optimistic to expect them to successfully navigate these intricacies.[171] By the criterion of simplifying a law that had become confusing, the MCA has manifestly failed.[172]

The Act faces a more fundamental challenge than the observation that it has failed to simplify the law. Article 12(2) of the UN Convention on the Right of Persons with Disabilities ('the UNCRPD') requires state parties to 'recognize that persons with disabilities enjoy legal capacity on an equal basis with others in all aspects of life'.[173] The UN Committee that oversees the Convention has argued that this is incompatible with any system in which a decision-maker 'can be appointed by someone other than the person concerned' to make decisions in the person's best interests rather than according to their 'will and preferences'.[174] This interpretation is controversial,[175] but it is premised on a different account of the relationship between autonomy and mental capacity to any found in domestic law. The survival account holds that those without mental capacity might retain such autonomy that their wishes determine their best interests, but the Committee may be committed to a universal account of autonomy. In this, everyone, regardless of mental capacity, always retains enough autonomy to restrain the state from acting against their will. Decisions must be made on the 'best interpretation' of a person's 'will and preferences':[176] a standard that, at least formally, ascribes autonomy even to those who cannot communicate. Commentators have doubted how 'realistic' the universal account is,[177] and it is arguably motivated by analysis that incorrectly treats the gatekeeper account

[170] Paul Skowron, 'Evidence and Causation in Mental Capacity Assessments' (2014) 22 Medical Law Review 631.
[171] Mary Donnelly, 'Capacity Assessment under the Mental Capacity Act 2005: Delivering on the Functional Approach?' (2009) 29 Legal Studies 464.
[172] The House of Lords Select Committee repeatedly observes that professionals do not understand the MCA (n 143) 6.
[173] (n 15).
[174] General Comment (n 18) para 27. Domestic law faces other challenges from this part of the UNCRPD. This section only addresses the challenge it poses to domestic accounts of the relationship between autonomy and capacity. For an overview, see Peter Bartlett, 'The United Nations Convention on the Rights of Persons with Disabilities and Mental Health Law' (2012) 75 Modern Law Review 752.
[175] Martin (n 136).
[176] General Comment (n 18) para 21.
[177] John Dawson, 'A Realistic Approach to Assessing Mental Health Laws' Compliance with the UNCRPD' (2015) 40 International Journal of Law and Psychiatry 70.

as literally true.[178] Nevertheless, it captures a significant point. It echoes the affirmation found in Kant that all humans are 'endowed with freedom'[179] and in Article 1 of the Universal Declaration of Human Rights that 'all human beings …are endowed with reason and conscience'.[180] The central insight of the universal account is that to deny that someone is capable of choice, even only for one decision or in a relatively small way, is to deny they have something often considered essential to being a full human being. Given the long history of marginalisation faced by those with mental disabilities, this concern should be taken seriously. Nevertheless, the universal account makes the same error as the three accounts found in domestic law. It assumes that a 'single plain litmus-paper test'[181] can capture all the relevant features of the concept of autonomy.

The Mental Capacity Act was to some extent motivated by a similarly reductive account of the relationship between autonomy and capacity, the gatekeeper one;[182] but those ambitions of laying down a simple rule did not survive even a few years contact with reality. They did not deserve to. Issues of voluntariness, subjectivity, and the intelligibility and authenticity of a person's wishes are important moral concerns; and it is right that judges have taken them seriously. The issue of personhood, which the UNCRPD makes it almost impossible to avoid, is at least as important. It is, however, no more capable of articulating all the details of the complex moral dilemmas that arise across the country hundreds of times every day than the concept of mental capacity has been. Legal reforms that take personhood seriously may be welcome, but any reform attempting to remake the law around that one concept can be expected to fail.

6. Conclusion

Judges have told three simple but contradictory stories about the relationship between autonomy and mental capacity in the law of England and Wales. Sometimes, capacity is autonomy's gatekeeper: those with capacity are autonomous, and those without capacity

[178] General Comment (n 18) para 15: 'The functional approach attempts to assess mental capacity and deny legal capacity accordingly'.
[179] Immanuel Kant, *The Metaphysics of Morals* (first published 1797, Mary Gregor tr, Cambridge University Press 1996) 6:280-281. An account of Kant's position is given in Patrick Kain, 'Kant's Defense of Human Moral Status' (2009) 47 Journal of the History of Philosophy 59.
[180] Universal Declaration of Human Rights (adopted 10 December 1948) UNGA Res 217 A(III).
[181] Midgley (n 157) 184.
[182] Lord Falconer LC (n 31); Code (n 33) para 1.1-1.3.

are not. Sometimes, capacity is necessary for autonomy but not sufficient, for voluntariness is also required. Finally, sometimes autonomy survives incapacity, and the wishes of a person without capacity can reflect their ability to be self-determining. These three accounts coexist in the current law; so no story of evolution, in which one account comes to replace another, can be told. Similarly, judges will switch between accounts depending upon the facts that they face in a particular case, so no story of judicial factions is plausible. Instead, each account should be taken as a partial description of a consistent whole.

On a coherent account, a person will be regarded as autonomous if they have capacity about the matter in question and are not subject to undue influence or coercion *or* if they lack capacity about the matter but the character of their present wishes and feelings nevertheless determines their objective best interests. This account lacks the simple force of judicial statements such as 'autonomy survives incapacity' or 'the statute respects their autonomy', and such statements have their place. They are a way that a judge can show that the case before them is profoundly important not only to the person it concerns but also, because that person is part of society, to the law itself. Nevertheless, the overall relationship between autonomy and mental capacity is also important, for these cases go to the very heart of the relationship between state and citizen. To understand that relationship, statements largely intended to address the facts of particular cases must be put to one side. When this is done, the relationship that emerges may be considered good or bad, depending on the criteria that it is evaluated by. It is, however, undoubtedly complex. The purpose of this chapter has been to draw attention to this complexity. If this analysis has any direct implication for reform, then it is only to be suspicious of easy answers, no matter who offers them and what those answers might be.

Chapter Four

A practice-derived justification for actions against a person's expressed wishes

1. Introduction

This chapter draws together the last two; and suggests that, rather than being central to the justification of actions against someone's expressed wishes, dual model autonomy is only of marginal importance. Chapter two introduced dual models: those that attribute autonomy only to those desires that the person would endorse, or at least not be alienated from, on reflection. It then showed that, despite recent legal literature that relies on dual models, they cannot plausibly justify actions against someone's expressed wishes in a way that is value-neutral. Chapter three then examined the relationship between autonomy and mental capacity in the courts; and showed that, although the judiciary's rhetoric changes from case to case, a coherent account of that relationship can be constructed.

It is already apparent from these two chapters that philosophers and the courts tend to use the word 'autonomy' in different ways. Dual model theorists develop consistent, sparse accounts of what autonomy essentially is. They hope these accounts can, among other things, guide decisions about whether or not to respect a particular expressed wish. Judges, in contrast, are not orientated towards theoretical elegance. As the last chapter shows, they decide whether or not to respect a particular expressed wish using the concepts of capacity, voluntariness, subjectivity, intelligibility, and authenticity; and they use the word 'autonomy' more or less rhetorically, with little worry about coherence between cases.

Section two of this chapter shows that the difference between dual model theorists and the judiciary is more fundamental than this. Judges are primarily concerned with the person's relationship to the world. Dual theories, in contrast, are primarily concerned with the person's relationship to themselves. Most of the time, the basic subject matter of the two is simply not the same. This does not mean that dual models are entirely irrelevant to mental capacity law. Sometimes, as part of a best interests decision, judges are concerned with the person's relationship to themselves. This is, however, only part of such decisions; and, far

from justifying decisions against a person's expressed wishes, dual model reasoning can, at most, merely overturn a justification that has already been formulated on other grounds. Compared to the role granted to them by some recent legal literature,[1] dual models are relatively marginal.

Section three of this chapter treats legal practice as though it were a speaker making a serious philosophical point. In other words, section two shows that judges look to a person's relationship to the world, not their relationship to themselves; but section three takes seriously the possibility that the judges are right to do so, and this is where a justification for actions against a person's expressed wishes *should* be found. Using Mill's example of a person who is about to unwittingly cross a collapsing bridge, it develops such a justification, which has four stages: prospective harm, false belief, necessity, and proportionality. Unless all four are present, a person's expressed wishes should be respected. This should not be taken as an attempt to make it easy to justify actions against a person's expressed wishes. As subsection 3.5 shows, the four-stage justification only makes the difficulties, both moral and epistemic, loom larger.

2. Dual models in the law

This section draws together the analysis of dual models in chapter two and the analysis of the Court of Protection's rhetoric and practice in chapter three. Chapter two showed that dual models of autonomy cannot plausibly justify actions against a person's expressed wishes in a way that is neutral between different conceptions of the good. Chapter three showed that judges in England and Wales use the word 'autonomy' inconsistently, but that their practice across a range of cases nevertheless implies a coherent account of when a person's expressed wishes must be respected. This section shows that this coherent account is not a dual model, although it contains one in a subsidiary role. It does this by re-examining, in turn, the three components of the coherent account: the gatekeeper, survival and insufficiency accounts.

2.1 The unreflective gatekeeper

The gatekeeper account is the view that a person must be treated as autonomous if they have mental capacity and need not be if they do not have mental capacity. As discussed in

[1] Chapter two, section 3.

the last chapter, it is an oversimplification: a person with capacity may be treated as lacking autonomy by reason of coercion or undue influence; and a person without capacity may be treated as having autonomy by reason of subjectivity, authenticity, or intelligibility. Nevertheless, the gatekeeper account contains an important element of truth. In law, the expressed wishes of a person with capacity will only be overridden in the presence of coercion or undue influence, and the expressed views of a person without capacity will only be considered overriding in the presence of subjectivity, authenticity, or intelligibility. Capacity is still of central importance; and, as the concept of capacity bears a superficial resemblance to dual model autonomy, this can give the impression that dual models are of similar central importance. As this section demonstrates, however, this apparent resemblance between mental capacity and dual models does not survive scrutiny.

Section 3(1) of the Mental Capacity Act ('the MCA'), provides that, for its purposes, a person is unable to make a decision if they are unable to understand, retain, use, and weigh the information relevant to that decision, or if they are unable to communicate their decision. Capacity in this sense is a mental competence. It is an ability to something. Dual models of autonomy also contain a mental competence, one often simply called 'competence'.[2] This competence is, not, however the same as mental capacity. It is, as in Christman's theory, competence to (1) form intentions based on a motivating characteristic such as a desire or value, and (2) critically reflect on that motivating characteristic.[3] There is a similarity between this and mental capacity, but also two important differences. The similarity is that both 'competence' and 'mental capacity' require that the person be able to form an effective intention. The first important difference is that mental capacity almost always requires an accurate understanding of the world but that reflective competence does not. The second difference is that reflective competence requires an ability to critically analyse the motivating desire but that mental capacity seldom does. In other words, mental capacity is about a person's relationship to the world, but reflective competence is about a person's relationship to their own desires and values.

A specific example makes the difference clearer: comparing reflective competence about a particular sexual desire to mental capacity to consent to sex. Once again, Christman's

[2] John Christman, *The Politics of Persons: Individual Autonomy and Socio-Historical Selves* (Cambridge University Press 2009) 155.
[3] *Ibid*.

model, because it is precise, provides a good example of a dual theory. Applying his model to a particular desire for sex, a person is competent only if they can form an intention to act on the desire and they can critically reflect on both it and any other relevant motivating characteristics, for instance a desire to please the other person.[4] This critical reflection must be effective. In other words, it must be able to result in either the person being alienated from the desire, strongly rejecting it, or in them endorsing the desire.[5] Reflective competence about a particular sexual desire therefore has three components: the desire itself; an ability to form an intention to act based on that desire; and an ability to evaluate that desire in a way that results in either alienation or non-alienation.

Mental capacity to consent to sex has been extensively litigated, so its criteria can be precisely specified.[6] As the MCA requires, the focus of the test is on the 'information relevant to the decision',[7] which may include the reasonably foreseeable consequences of any decision or a failure to decide.[8] As with reflective competence, it has three components, but they are three very different components. A person has mental capacity to consent to sex only if they are able to understand 'i) the mechanics of the act; and ii) that there are health risks involved; and iii) that he or she has a choice and can refuse'.[9] The first two of these criteria are obviously nothing to do with reflective competence. They are not about the person's desire, ability to form an intention based on that desire, or ability to evaluate that desire. They are about the objective question of what sex is and what some of its consequences might be. The third criteria, that the person knows that they have 'a choice and can refuse', may, however, seem closer to reflective competence. It is not. This criterion is not about the person's ability to form a second-order response to their own sexual desire. It is about another objective fact, albeit one about society. This is clear when the criterion is applied:

> Although she found sex enjoyable and comfortable she had no idea that she had a choice and could refuse. Indeed the attitude of [her partner], based, as he told

[4] *Ibid*.
[5] *Ibid* 146. As discussed in chapter two, section 4, autonomy in Christman's model requires both this competence and 'authenticity', which is non-alienation on hypothetical reflection.
[6] If a person is found to lack this mental capacity, a decision-maker cannot consent to sex on their behalf: MCA s27(1)(b).
[7] *Ibid* s3(1).
[8] *Ibid* s3(4).
[9] *London Borough of Tower Hamlets v TB* [2014] EWCOP 53, [2015] 2 FCR 264 [41] (Mostyn J).

> me on his culture and religion, was that he had a right to seek sex from her and that it was her duty to submit. The evidence clearly showed that TB comprehensively failed the second and third criteria as formulated by me.[10]

For TB to understand consent required her to understand that, despite her partner's expectations, she had 'no duty to submit'. This is an important fact about society. It is not a fact about TB's reflective attitude to her own desires. In another case, Parker J says that the 'ability to understand the concept of and the necessity of one's own consent is fundamental to having capacity' for sex.[11] Once again, it is easy to be misled. Read without care, this can look something like reflection: 'understand …one's own consent'. It is nothing of the type. The full phrase 'the concept of and the necessity of one's own consent' makes it clear that Parker J is referring to the person understanding what consent is, and that without their own consent sex cannot legally occur.[12] This is not the person's reflective attitude to their own consent.

It is clear from both the statute and the case law on capacity to consent to sex that reflective competence is not a part of a person's 'understanding' under section 3(1)(a) of the MCA. If it does not fall under that part of the functional test, however, then it is no part of the test at all. The only other possible candidate, section 3(1)(c), reads 'to use or weigh *that information* as part of the process of making the decision' (emphasis added), referring back to section 3(1)(a). In other words, it does not allow the inclusion of information that has not been considered relevant in the 'understanding' phase of assessment, and any attempt to introduce new information in the 'using or weighing' stage merely represents a failure to adequately respect the statutory language.

This fundamental difference between mental capacity and reflective competence is not limited to cases about capacity to consent to sex. Indeed, it exists even when judges are taking some care to assess a person's understanding of their own mental states. For instance, consider the following assessment of whether a young woman with Asperger Syndrome, learning disabilities, and a history of risk-taking has capacity to decide whether to live independently or with her mother:

---

[10] *Ibid* [45].
[11] *London Borough of Southwark v KA* [2016] EWCOP 20, [2016] COPLR 461 [53].
[12] *Ibid* [53]-[56].

> She recognised that she would reasonably expect to be allocated a flat, and was able to distinguish between the 'good' areas and 'bad' areas of town in which to live. I am (perhaps unlike Dr. Rippon) sufficiently persuaded that Z recognises at a material level the benefit of third party support in the event that she is to live on her own. She showed insight into the possible loneliness of living independently; she felt that one of the downsides of leaving home is that she will lose the benefit of having her mother's "shoulder to cry on" when things are getting her down. She has an outline knowledge of her financial circumstances, and currently appears able to perform basic budgeting. She seems aware that her life is easy now, as all the bills are paid, and she is cared for; I felt that she recognised that she would be giving these comforts up if she were to move.[13]

Mr Justice Cobb considers Z's understanding of a number of facts about the world: what kind of residence would be provided; what different neighbourhoods might be like; basic finances; and social support. He also considers her understanding of the possible consequences of independent living on her own emotional state. Even this, however, is not an assessment of Z's reflective competence. An assessment of her reflective competence with regard to her desire to live independently would assess the following: that desire's presence; her ability to form an intention based on it; and her ability to critically reflect on that desire. Cobb J takes the first two of these as obvious, and does not mention the third. Instead of her attitude to her own desire, he is concerned, as the Act demands, with whether she can understand, use, and weigh a plausible emotional consequence of following that desire. This is certainly a mental competency that relates to the person's own mind. It is not the same mental competence as appears in dual theories. A dual theory would be concerned with how Z feels, or would feel on reflection, about her own desire to live independently. Cobb J is concerned about whether she understands that living away from her mother might make her feel lonely.

2.2 'Insight' is not reflective competence

It is not enough to show that the law does not require reflective competence for mental capacity. There is evidence that capacity assessments often diverge from the statutory

[13] *WBC (Local Authority) v C* [2016] EWCOP 4 [61] (Cobb J).

language,[14] so reflective competence might nevertheless be demanded in practice. There is little evidence that this happens. The most likely candidate for subverting the Act in this way, unanalysed use of the concept of insight, remains closer to the statutory language than it does to a dual theory. It must be admitted, however, that 'insight' does feature heavily in capacity assessments. Although the word does not appear in the Act, a recent study found about a third of health and welfare cases in the Court of Protection that discussed the person's capacity made reference to their 'lack of insight'.[15] This is, as Case argues, worrying. Insight is poorly defined and reliance on the concept can obscure the application of the statutory criteria for incapacity.[16] Insight also seems like a concept closely related to critical reflection: practitioners and judges repeatedly talk about a person's insight into their own 'condition' or 'needs'.[17] Once again, however, it is important not to assume that all mental competencies are the same.

The criteria for dual model reflective competency are a desire (or other motivating characteristic), the ability to form an intention based on that desire, and the ability to reflect on that desire in a way that results in alienation or endorsement. The examples in the emerging 'insight' literature do not fit these criteria. In the Court of Protection, 'insight' has been used to refer to whether the person believes their diagnosis, or understands their broader 'condition', 'problems', 'abilities', care needs, care package, salient risks, and even the care needs of others.[18] Similarly, interviews with practitioners have found insight being used to refer to a person's understanding of their own needs, of 'problems', or of potential 'consequences' with their decision.[19] Given the width of this subject matter, it is hardly surprising that there is concern that 'insight' is sometimes used by professionals to turn a disagreement with the person into a lack of capacity that gives the professional freedom to determine the person's best interests.[20] None of these uses, however, are the same as

[14] Val Williams and others, *Making Best Interests Decisions: People and Processes* (Mental Health Foundation 2012) 50-52.
[15] Paula Case, 'Dangerous Liaisons? Psychiatry and Law in the Court of Protection — Expert Discourses of 'Insight' (and 'Compliance')' (2016) 24 Medical Law Review 360, 369.
[16] *Ibid* 379.
[17] *Ibid* 370, 373; Williams (n 14) 52-58, 72.
[18] Case (n 15) 369-373.
[19] Williams (n 14) 53-58; Charlotte Emmett and others, 'Homeward Bound or Bound for a Home? Assessing the Capacity of Dementia Patients to Make Decisions about Hospital Discharge: Comparing Practice with Legal Standards' (2013) 36 International Journal of Law and Psychiatry 73.
[20] Neil Allen, 'Is Capacity "In Sight"?'(2009) 19 Journal of Mental Health Law 165; Williams (n 14); Emmett (n 19); Case (n 15).

reflective competence. None of them are about the person's ability to reflect on their own desire. Many are simply about the person's understanding of the world: care packages, risks, the needs of others. Even those uses of 'insight' that are more intimate are still about the person's relationship to the world: what their needs, problems, and abilities as a person in the world are; and whether a particular diagnosis is an objectively true description of them. This, unlike dual models of autonomy, is not an assessment of the person's relationship to their own desire.

This difference between dual models and the more intimate forms of 'insight' can be made clearer by example. Consider two people with a diagnosis of depression, Winston and Ludwig. Winston denies that he has depression. He lacks 'insight' into his condition. This does not mean that he is necessarily alienated from his depressive feelings and desires. Indeed, because he denies that these feelings are symptomatic of depression he is probably less likely to reject them. Ludwig, in contrast, admits that he has depression. He has 'insight'. At the same time he despises the feelings associated with depression. He is alienated from them. In other words, Winston lacks insight but has dual model autonomy, and Ludwig has insight and does not have dual model autonomy. There is no necessary correlation between the two. The reason for this simple. Dual models are about a person's relationship to themselves as an intentional agent in the world: their assessment of their own intentional states, such as desires. 'Insight', in contrast, is about a person's relationship to themselves as an object in the world, whether they fit third-person descriptions such as 'a person with depression' or 'a person who needs support'.[21] In everyday life, people understand themselves and one another in both of these ways, as both agents and objects in the world, but there is no way of reducing one mode of explanation to the other.[22] They are simply not the same.

Insight, like mental capacity, is about a person's relationship to the world. Even when it does seem to be about the person's relationship to themselves it is to themselves as an object

[21] Intentionality is a huge area of philosophy, but the basic concept is that intentional states are, in Searle's words, 'directed at objects and states of affairs in the world'. For instance, a desire is a desire *for* something, and a belief is a belief *about* something. Careful distinctions are required when dealing with psychiatric diagnoses. If I am depressed about losing my job, then my depression is an intentional state: it is directed at the loss of my job. In contrast, to say that I 'have depression' is to draw attention to a global feature of my mental states that is not, itself, necessarily intentional. John R Searle, 'What Is an Intentional State?' (1979) LXXXVIII Mind 74.

[22] Mary Midgely, *The Ethical Primate: Humans, Freedom and Morality* (Routledge 1994) ch 6.

in the world, not to themselves as an agent. Reflective competence is about a person's relationship to their own agency. These two mental competencies are not the same. If that is so, however, then dual model autonomy can form no fundamental part of the justification for treating those without capacity or insight differently to those with capacity or insight. As chapter two found that dual models cannot plausibly justify actions against a person's expressed wishes, this finding may be reassuring. Possibly, practice has a better justification. Statute, case law, and practitioner interviews all converge on the person's relationship to the world. Section three takes this seriously as an alternate base of justification. This subsection has not, however, shown that dual models are entirely irrelevant. They play a role in best interests assessments.

2.3 A reflective brake

Dual theories do appear in mental capacity law, but in a counterintuitive place. The last chapter discussed cases that demonstrate a 'survival' account of the relationship between autonomy and capacity, ones in which a person without capacity is nevertheless treated as being 'autonomous' (in the legal sense of that term). It also noted that 'authenticity' concerns in such cases seemed similar to philosophical dual models. This use of dual models is, however, profoundly different to the idea that a person's lack of autonomy can justify actions against their expressed wishes. Instead, actions against their expressed wishes are already presumptively justified by their lack of capacity, which is not reducible to dual model autonomy; but this presumptive justification can sometimes be overturned by the presence of something similar to dual model autonomy. In other words, the absence of dual model autonomy does not justify actions against someone's expressed wishes, but its presence may sometimes counter a separate justification for acting against a person's expressed wishes.

As discussed in the last chapter, the survival account of the relationship between autonomy and capacity is a heterogeneous mixture of different factors. The structural factors, the person's closeness to the 'borderline' of capacity and whether their wishes can be accommodated within their overall best interests, merely use the Act's existing categories; so they do not offer an avenue for dual theories to be introduced. The more substantive factors – subjectivity, intelligibility, and authenticity – demand closer attention. Intelligibility can be dealt with shortly, for it follows the same pattern found with capacity

in subsection 2.1. It is 'the extent to which P's wishes and feelings are, or are not, rational, sensible, responsible and pragmatically capable of sensible implementation in the particular circumstances'.[23] As with mental capacity, this is not about the person's relationship to their own desire, but about the relationship between that desire and the world. As such, it is not dual model autonomy. Nor is the second substantive factor from the survival view, subjectivity, the same as dual model autonomy. This was given in *ITW v Z* as 'the possible impact on P of knowledge that her wishes and feelings are not being given effect to'.[24] This, too, is about the relationship between the world and the person's desires, not about the person's relationship to their own desire. Although this strand of cases is about the person as a subject, about considering 'matters from the patient's point of view',[25] it is about the relationship between that subject and the world, not the relationship between that subject and themselves.

The final substantive factor, authenticity or 'the strength and consistency of the views being expressed by P',[26] is more interesting from the perspective of dual model autonomy. Only one type of dual model seems to have so far appeared in these cases, one based on the person's real reflection.[27] In comparing these cases with dual models, chapter two's central questions – what is the standard of reflection and how is it measured? – are useful for teasing out the details of the court's reasoning. The case showing the strongest evidence of something like a dual model is *Wye Valley NHS Trust v B*.[28] As discussed in chapter three, the question in this case was whether it was in the best interests of a man with schizophrenia to have his leg, which had a life-threatening infection, amputated against his wishes.[29] Reflection plays a prominent role in the case, and in the decision that amputation was not in Mr B's best interests. Mr Justice Peter Jackson noted that 'the wishes and feelings, beliefs and values of people with a mental disability are as important *to them* as they are to anyone

[23] *ITW v Z* [2009] EWHC 2525 (Fam), [2011] 1 WLR 344 [35] (Munby J).
[24] *Ibid.*
[25] *Aintree University Hospitals NHS Foundation Trust v James,* [2013] UKSC 67, [2013] 3 WLR 1299 [45] (Hale L).
[26] *ITW* (n 23) [35].
[27] An extended consideration of what a man in a minimally conscious state would want occurs in *Briggs v Briggs* [2016] EWCOP 53, [2017] 4 WLR 37 [101]-[115], [129]. This is an exercise in hypothetical reasoning, but not hypothetical reflection. The former asks what a person would want, but the latter ask what the person's attitude to their own wants would be.
[28] [2015] EWCOP 60, [2015] COPLR 843.
[29] *Ibid* [3].

else',[30] 'Mr B's religious sentiments are extremely important to him',[31] and 'he is a proud man who sees no reason to prefer the views of others to his own'.[32] In contrast with the cases discussed so far in this chapter, these statements are about Mr B's relationship to his own values and desires, not to the relationship between those values and desires and the world. They deal with the right subject matter to represent a dual model. They are also clearly about real reflection: Mr B's religious sentiments 'are', not 'would be', important to him; and he 'sees' no reason to prefer the views of others, not 'would see' no reason. The judge even includes a selection of statements made by Mr B in which he talks about his own desires and lack of fear.[33] It seems clear that it is the presence of real reflection, not hypothetical reflection or mere competence, which has countered the justification for decisions against Mr B's expressed wishes.

The standard of reflection to which Mr B has been held is harder to define, although it is clear that Peter Jackson J believes he meets the standard: 'His religious beliefs are deeply meaningful to him and do not deserve to be described as delusions: they are his faith and they are an intrinsic part of who he is'.[34] This is clearly some sort of reflection over time. The judge emphasises that they are 'of such long standing that they are an inextricable part of the person that he is';[35] and, immediately before quoting Mr B's statements about his own wishes, stresses that he 'has consistently said this over the entire period that amputation has been under discussion'. Something like Christman's requirement that critical reflection be sustained over a variety of conditions seems to be operating here.[36] It is similarly clear that Peter Jackson J is not demanding entirely accurate reflection, for he states that Mr B 'does not understand the reality of his injury'.[37] This leaves two possible standards of reflection that Mr B could be being held to: either bare reflection over time, in which any sustained reflection over a variety of conditions is adequate, or reflection that is both sustained across time and held to some other substantive standards, while falling short of complete accuracy. The evidence in the judgment is ambiguous. No clear substantive

[30] *Ibid* [11] (emphasis added).
[31] *Ibid* [14].
[32] *Ibid* [43].
[33] *Ibid* [37(1)].
[34] *Ibid* [43].
[35] *Ibid* [13].
[36] Christman (n 2) 155.
[37] *Wye Valley* (n 28) [34(2)].

standards are set, but the judge draws a careful parallel between Mr B's 'religious sentiments' and 'religious belief' more generally.[38] This suggests that because Mr B's reflection resulted in him refusing treatment for reasons that, if unusual, do occur in people with no psychiatric problems it was easier to endorse that reflection.[39] Similarly, Mr B spoke cogently about death, saying 'I'm not afraid of dying' and 'I'm not afraid of death. I don't want interference. Even if I'm going to die, I don't want the operation'.[40] If he had refused treatment, but entirely denied that he risked death would his wishes have carried as much weight? It is impossible to know for sure, but the suspicion that they may not have done perhaps implies that Mr B's reflection is being held to some further standards, if only epistemic standards about his responsiveness to evidence. In the same manner, if there was evidence that Mr B's views were the result of coercion or undue influence, as in *Re T (Adult: Refusal of Treatment)*,[41] it seems at least plausible to suggest that the case may have been decided differently. Taken together, these points are not sufficient to show that Mr B was being held to a higher standard than bare reflection over time, but they are enough to suggest that it is possible.

It is important to proceed methodically when unravelling what *Wye Valley* might show about justifying actions against a person's expressed wishes. Although Mr B's attitudes to his own wishes are prominent in the case, they are not where Peter Jackson J started. Instead, he began, as required by the Act, by determining the question of capacity. In doing so, his attention was on the relationship between Mr B's wishes and the world, not the relationship between Mr B and his wishes: 'there are limitations in Mr B's ability to understand the information about *his damaged foot* and a clear inability to weigh the *relevant medical evidence*'.[42] This disconnection between Mr B's reasoning and the world led to a finding that he lacked capacity,[43] and it is this that mandated a best interests decision that may have overridden his expressed wishes. In other words, the Court's first finding was that an action against his wishes might be legitimate, and the justification for that finding had nothing to do with dual model autonomy. If it had found that he had capacity, then

[38] *Ibid* [14]-[15].
[39] Hayden J notes that 'It is notable that in the Wye Valley case, B's beliefs, though in part delusional, were connected to a profound religious belief': *NHS Foundation Trust v QZ* [2017] EWCOP 11 [31].
[40] *Wye Valley* (n 28) [37(1)].
[41] [1992] EWCA Civ 18, [1993] Fam 95, 116.
[42] *Wye Valley* (n 28) [35] (emphasis added).
[43] *Ibid.*

unless coercion or undue influence were shown, no action against his expressed wishes would have been legitimate, regardless of how alienated he may have felt about those wishes on reflection. When, however, a lack of capacity was found, then an action against his wishes was nevertheless not found to be proportionate, and dual model autonomy was an overriding factor in that conclusion.[44] In law, a mere lack of dual model autonomy does not justify actions against a person's expressed wishes. The presence of real reflection, however, may delegitimise actions against a person's expressed wishes that would otherwise be permitted due to their failure to understand, use, and weigh relevant information about the world. Sometimes, and only sometimes,[45] a person's reflective endorsement of their own desires may act as a brake on the ability of others to act against those wishes; but this brake is only relevant when such actions would otherwise be legitimated on other grounds. Chapter two concluded that dual models cannot, by themselves, legitimate actions against a person's expressed wishes. In contrast, the analysis here shows that, in law, they do not. Their role in justification is marginal, a point discussed in more detail in section three. First, however, there is another type of case that was discussed in the last chapter.

2.4 Dual models and the insufficiency account

Chapter three drew attention to cases in which a person with mental capacity was nevertheless treated as lacking autonomy, in the sense that their expressed wishes were not treated as final. It called these lines of case law 'the insufficiency account', drawing a parallel between the equitable doctrine of undue influence and some uses of the inherent jurisdiction of the High Court. This subsection shows that these cases present at least three possible relationships between undue influence and a person's wishes: coercion can stop someone acting on their desires; it can stop someone from expressing their desires; or it can change the content of their desires. In the cases, these three categories sometimes coexist,

[44] An *obiter* comment in *Briggs* (n 27) perhaps shows the other side of this coin: 'it is not uncommon that very understandable expressions of present wishes and feelings "I want to go home" would not be made if P was able to weigh the existing competing factors by reference to P's beliefs and values' [60(iii)] (Charles J). In other words, if P has not reflected, their wishes may be accorded less weight.

[45] Another case where the person's reflection on their own desires carried at least some weight is *A NHS Foundation Trust v Ms X* [2014] EWCOP 35, [2015] COPLR 11, particularly at [51]. It is, however, striking that reflection is not usually a major issue even in cases that cite *Wye Valley*. For example, in *SAD v SED* [2017] EWCOP 3, Glentworth DJ's emphasis is on the realism of SED's views more than her reflective endorsement of them ([28]-[30]).

and judges sometimes speak so impressionistically that it is hard to be sure what relationship between the undue influence and the person's expressed desire they believe to exist, but for the purposes of analysis it is important to keep them separate. Once this is done, it can be shown that none of these categories are reducible to dual model autonomy, and there is little evidence of anything like dual model autonomy operating in these cases. Despite this, the open ended nature of the inherent jurisdiction, which has been described as probably having 'no theoretical limit',[46] means that it cannot be said that dual models will never be relevant in the future.

The first relationship between undue influence and person's desires is relatively straightforward. There are situations in which a person can formulate desires, but their ability to act on them is restricted by overt coercion. For instance, in *Al-Jeffery*,[47] which concerned a young woman who had been locked in a room in her father's flat in Saudi Arabia for over a year,[48] the woman was clear that 'she would like to leave the country but her family are not allowing her to leave'.[49] Similarly, in *SA*, there was concern 'that SA might be about to be taken by her family to Pakistan to be married there to some unknown person *contrary to her wishes*'.[50] There is no trace of dual model autonomy in these passages. The person has formed a desire and is not alienated from it. It is simply that their ability to successfully act on the desire has been restricted by others. As with capacity assessments, the issue here is not the person's relationship to their own motivating characteristics. The issue is the world, in the form of other people, restricting someone's ability to act as they wish. Once again, this is simply a different subject matter to whether or not the person would be alienated if they reflected on their own desires.

The second possible relationship between undue influence and person's desires is similarly unrelated to the subject matter of dual models. In some cases, a person is believed to have clearly formed wishes that they are afraid to express because of the coercion or undue influence. Then, whether by silence or by being coerced into lying to third parties, they might express desires that they do not have. This was an issue in *Re T (Adult: Refusal of*

[46] *A Local Authority v SA* [2005] EWHC 2942 (Fam), [2006] 1 FLR 867 [45] (Munby J).
[47] *Al-Jeffery v Al-Jeffery* [2016] EWHC 2151 (Fam), [2016] Fam Law 1319.
[48] *Ibid* [19].
[49] *Ibid.*
[50] *Re SA* (n 46) [18] (emphasis added).

Treatment),[51] where Lord Donaldson MR stated that part of the 'real question' in such cases is 'does the patient really mean what he says or is he merely saying it for a quiet life, to satisfy someone else'?[52] Something similar appeared to be an issue in *London Borough of Redbridge v G*,[53] a case that was ultimately decided under the MCA. In that case, Ms Justice Russell was concerned that threats and deception by live-in carers had left a 94 year old woman 'subdued' and 'distrustful'.[54] Indeed, G 'asserted in a letter in the clearest terms that she was happy with C as her full time live-in carer and wanted no more involvement with social services, only to say later, at a safeguarding meeting, that the letter had been dictated by C'.[55] With regards to the original letter, the issue here is not, as in *Al-Jeffery*, of a person able to express their wishes but unable to act on them. Instead, it is an issue of someone who, due to undue influence, cannot even express their wishes. Even such situations, however, are not analogous to dual models of autonomy. Dual models examine the person's reflective attitude to their own desire: for instance, whether or not they feel alienated from it. With this second kind of coercion, however, the issue is whether a person's expressed desire matches their actual desire. Once again, this is not the same thing. It would assume too much to ask whether G in *Redbridge* felt alienated from her desire for social services to no longer be involved in her care. C's coercion calls into question the idea that G even felt the desire at all. Both dual models and this sort of coercion involve a mismatch; but this sort of coercion is about a mismatch between what a person desires and what they express, whereas dual models concern a mismatch between a desire and what the person would like to desire.

A third relationship between undue influence and the person's desire is also easily mistaken for something like dual model autonomy. Sometimes it is thought that a person's desires are themselves the result of coercion or deception. This, too, occurred in *Re T*. The second part of what Lord Donaldson called the 'real question' in such cases was whether 'the advice and persuasion to which he has been subjected is such that he can no longer think and decide for himself?'[56] It also seemed to be a factor in *Redbridge,* where reference

[51] (n 41).
[52] *Ibid* [31].
[53] [2014] EWCOP 17.
[54] *Ibid* [20].
[55] *Ibid* [80].
[56] *Re T* (n 41) [31].

was made to G's will changing,[57] ideas being drummed into her,[58] and the 'manipulation' of her wishes.[59] In these situations, the thought seems to be that if it was not for the coercion or deception, then the person's wishes would be different. In other words, a kind of hypothetical, counterfactual, reasoning is occurring: something of the type 'if *P* had not been coerced, then she would not have desired *X*'. It is easy to be hasty here, and to conclude that because counterfactuals about desire also occur in dual models this must be a dual model. It is not. This section has already noted that not all mental competencies are the same and not all mismatches of desires are the same. Similarly, not all counterfactuals about desire are the same. In cases concerned with undue influence and coercion, the question that the judge asks is whether the person would have had a particular desire if it was not for the coercion or deception.[60] A dual model, in contrast, would ask what the person's attitude to their own desire would be if they reflected upon it. The two do not coincide. A person might reflectively endorse a desire caused by coercion; and, conversely, a person may feel alienated from a desire that was not caused by coercion. As with capacity assessments, the undue influence cases are concerned with the connection between the world and the person's desires. Dual models, in contrast, are concerned with the relationship between the person and their own desires.

Although it has not been possible to find any evidence of a dual model in the inherent jurisdiction case law, its open-ended nature means that this could change in the future.[61] One possibility is that something analogous to the place of dual models in best interests decisions may occur. It seems at least possible for a court to find that a third party has applied undue influence to a person; but that the person's influenced decisions nevertheless coincide with their authentically held desires, and should be allowed to stand. For instance, if undue pressure was exerted on someone to donate to a genuine charity, but the cause was one that the person had long been sympathetic to and the amount they gave was within their usual range for such donations, then it seems possible that a judge could, with appropriate injunctions, stop the undue influence but allow the person's existing donations to stand. For now, this possibility remains speculative; but, if it occurs, then dual theories

[57] *Redbridge* (n 53) [20].
[58] *Ibid* [28].
[59] *Ibid* [58].
[60] *Re T* (n 41) [31], [32], [37(6)]; See also *DL v A Local Authority* [2012] EWCA Civ 253, [2013] Fam 1 [57]-[63].
[61] Furthermore, not every case is reported, so dual models may already be influencing other cases.

would be playing a role analogous to that they play within the MCA case law. Rather than justifying an action against the person's expressed wishes, they would be overriding decisions against the person's expressed wishes that had been justified on other grounds.

2.5 The place of dual models in the current law

Chapter two of this thesis suggested that actions against a person's expressed wishes in their best interests could not be justified by dual models alone. Chapter three developed an account of the actual practice of the judiciary. This section has developed that account to show that no attempt is made in law to use dual models as a sole basis of justification. Instead, dual models sometimes play a subsidiary role, preventing decisions against a person's expressed wishes when they would otherwise be justified.

The crux of current legal justifications for actions against a person's expressed wishes is not in the person's relationship with themselves. It is in the person's relationship with the world. Importantly, this is a relationship with two possible 'directions of fit':[62] person to world and world to person. Problems in either direction may justify the conclusion that the person should not be treated as 'autonomous' (in the legal sense of that term). As discussed in subsections 2.1 and 2.2, if there is a relevant problem in the person's understanding of the world, and they cannot be supported to change that understanding, then they may be found to lack capacity; and so actions against their expressed wishes may be presumptively justified. In these cases, a problem has been found to exist in the person's relationship to the world. As discussed in subsection 2.4, however, sometimes coercion or undue influence may stop a person from feeling able to express their desires or cause them to have desires that they would not otherwise have had. In these cases, a problem has been found to exist in part of the social world's relationship to the person. Neither direction of this relationship, person to world nor world to person, can be usefully reduced to dual model autonomy, which concerns a person's relationship to aspects of themselves.

Dual model autonomy does, however, have a subsidiary role in legal justifications of actions against a person's expressed wishes. When a decision against a person's wishes is presumptively justified because of a problem in their understanding of the world, then the

[62] John Searle, 'A Taxonomy of Illocutionary Acts,' in K Gunderson (ed), *Language, Mind and Knowledge* (University of Minnesota Press 1975) 344, developing a distinction made in GEM Anscombe, *Intention* (first published 1957, 2nd ed, Harvard University Press 2000) §32.

presence of real reflection may overturn that presumption. In legal terms, when a person is found to lack capacity, then sometimes it is not in their best interests to act against their desires simply because of how authentically those desires are held. It is possible that something similar might occur in some cases of coercion, although there is no obvious evidence of this in the recorded cases.

In one respect, these findings are reassuring. Chapter two showed that dual model autonomy alone cannot plausibly justify actions against a person's expressed wishes, so it is a good thing that it does not. This negative point, however, is not the entire story. The work in this section also allows the development of a positive account of the justification for actions against a person's expressed wishes.

3. How to justify decisions against a person's expressed wishes

Iris Murdoch once complained that 'for the hard idea of truth we have substituted a facile idea of sincerity'.[63] Her distinction is apt. Dual theories, by locating autonomy in the person's relationship to their own desires, demand only sincerity. The law, locating it in the person's relationship to the world, still demands truth. This section proceeds on the basis that this commitment to truth is justifiable. Ultimately, autonomy in the dual model sense does not matter without reality. As Taylor says, 'unless some options are more significant than others, the very idea of self-choice falls into triviality and hence incoherence'.[64] Why should it matter if I am alienated from or endorse my choices, if none of my choices matter? Some commitment to truth seems necessary; and, if so, the law might not just justify actions against a person's expresses wishes on a different basis to dual theories. It might justify such actions on a better basis. This section takes this idea seriously, and abstracts from practice a justification for actions against a person's expressed wishes grounded in their relationship to the world. Briefly, it holds that there are four necessary conditions for actions against a person's expressed wishes: prospective harm; false belief; necessity; and proportionality. Unless all four are present, no action against the person's expressed wishes is legitimate.

[63] Iris Murdoch, 'Against Dryness' (1961) 16 Encounter 16.
[64] Charles Taylor, *The Ethics of Authenticity* (Harvard University Press 1991) 39.

Chapter two traced dual models to Mill's *On Liberty*.[65] That same book, however, also contains an example that accords better with the practice of the courts:

> If either a public officer or any one else saw a person attempting to cross a bridge which had been ascertained to be unsafe, and there were no time to warn him of his danger, they might seize him and turn him back, without any real infringement of his liberty[66]

Mill continues with a dual model justification: '…for liberty consists in doing what one desires, and he does not desire to fall into the river'.[67] In other words, he posits a hypothetical deep self with desires different to the empirical person. As a dual model, this falls into all the problems with setting a standard of reflection discussed in chapter two. Indeed, this dual model is particularly unconvincing. In the example, there is no time to warn the person; but, without argument, Mill expands the scope of this principle to children, the 'delirious', and anyone 'in some state of excitement or absorption incompatible with the full use of the reflecting faculty'.[68] These cases are not like the example. Due to the shortage of time in the example it must be presumed that the person does not wish to die,[69] but in the other cases there is no need to presume anything about the person's desires. It is possible to talk to the person. Furthermore, Mill's justification even fails for the example itself. The person is expressing a desire, to cross the bridge, in their behaviour. They are *trying to cross*. Mill also imputes a desire, to not die, to them. This imputation has problems: after all, people do sometimes want to die. Even if it is accepted, however, Mill's justification fails. He relies on 'what one desires',[70] as if the intervener, by seizing the person, were just giving effect to a single, unitary desire. That is not what Mill has to justify. He has to justify giving effect to one desire, to live, over another, to walk onto the bridge. This he does not do. His dual model fails to do the work demanded of it. Despite this, his example is valuable. It is a paradigmatic case in which action against the person's

[65] John Stuart Mill, *On Liberty* (first published 1859, David Bromwich and George Kateb eds, Yale University Press 2003).
[66] *Ibid* 158.
[67] *Ibid.*
[68] *Ibid.*
[69] Jillian Craigie, 'Capacity, Value Neutrality and the Ability to Consider the Future' (2013) 9 International Journal of Law in Context 4, 12.
[70] Mill (n 65) 158.

expressed wishes would generally be considered acceptable. For this reason, reflection on it can help to develop an account of the conditions necessary for such action.

3.1 Risk of Harm

The most obvious condition in the example is prospective harm. Unless the bystander honestly believes that the bridge is unsafe, then they do not have any obvious justification for acting against the person's expressed wishes. There is a strong connection here with professional practice. As Glover-Thomas notes, 'nebulous and fluid' conceptions of risk guide professional behaviour in clinical mental health settings;[71] and, as Fanning has argued, consideration of risk remains central to determining whether a person is eligible to be deprived of their liberty under Schedule A1 of the MCA.[72] More widely, perception of risk is both an extremely common trigger for mental capacity assessments and a 'strong guide' to best interests decisions when someone is found to lack capacity.[73] Indeed, it has been suggested that ability to perceive risks is intrinsic to the concept of mental capacity;[74] and the statutory test, which requires the person be able to understand, use, and weigh the 'reasonably foreseeable consequences' of a decision,[75] will certainly often require some understanding of objective risks. Further evidence that risk of harm plays an important role in the current law is indirect. A suggested binary between 'autonomy' and 'protection' permeates much of the legal literature,[76] and 'protection' obviously implies some risk that the person is being protected from.

It cannot be simply concluded that just because practice relies on ideas of risk it should do so. Nevertheless, making risk irrelevant would result in more actions against people's expressed wish, not fewer. Although considerations of risk are a weak and malleable

[71] Nicola Glover-Thomas, 'The Age of Risk: Risk Perception and Determination Following the Mental Health Act 2007' (2011) 19 Medical Law Review 581, 586.
[72] John Fanning, 'Continuities of Risk in the Era of the Mental Capacity Act' (2016) 24 Medical Law Review 415.
[73] Williams (n 14), §2.3, §8.1.
[74] Ian Wilks, 'The Debate over Risk-Related Standards of Competence' (1997) 11 Bioethics 5.
[75] MCA ss3(1), 3(4).
[76] Some limits of this binary are explored in chapter seven. Examples of its pervasiveness (there are many more): James W Ellis, 'Decisions by and for People with Mental Retardation: Balancing Considerations of Autonomy and Protection' (1992) 37 Villanova Law Review 1779; Generva Richardson, 'Balancing Autonomy and Risk: A Failure of Nerve in England and Wales?' (2007) 30 International Journal of Law and Psychiatry 71; MC Dunn and others, 'To Empower or to Protect? Constructing the 'Vulnerable Adult' in English Law and Public Policy' (2008) 28 Legal Studies 234.

constraint on actions against a person's expressed wishes,[77] if such actions are only permitted where there is a real risk of harm, that is still a constraint. In other words, there may be problems with professionals considering risk in practice, but the problem is not that they think about risk at all. It is that they only consider risk, and not the other conditions necessary to justify an action against someone's expressed wishes.[78]

3.2 False belief

A second necessary condition in Mill's bridge example is false belief. For Mill, if the person knows that the bridge is unsafe, then 'seizing' them is unjustified.[79] As discussed in section two, mental capacity laws are similarly about the person's relationship to aspects of the world. Indeed, their emphasis on understanding suggests a second criterion for acting against a person's wishes: the prospective harm must be caused by a false belief that the person holds.[80] In this context, an absence of relevant true belief can count as false belief. The person in Mill's example need not actively think 'this bridge is safe'. It is enough if they fail to think 'this bridge is unsafe'. On false belief, the law is more defensible than dual models are. As Killmister points out, autonomy theorists have neglected the question of whether or not a person's beliefs about the world are true; but unless a person's relevant beliefs about the world are true then the relationship between what they intend to do and what they actually do entirely breaks down.[81] A person might be autonomous in the narrow sense of reflectively endorsing their own desire; but if that endorsed desire is based on beliefs that are demonstrably false, it is not obvious why it should always be respected when acting on it will result in predictable harm to the person.[82] It is hard to see why the person in Mill's example reflectively endorsing their desire to get to the opposite bank of

[77] Glover-Thomas (n 71); Williams (n 14).
[78] Williams (n 14) §3.2.
[79] Mill (n 65) 158.
[80] In cases of an inability to 'use or weigh' information this can be taken as a second-order false belief about the relative importance of, possibly true, first-order beliefs that the person holds.
[81] Suzy Killmister, 'Autonomy and False Beliefs' (2012) 164 Philosophical Studies 513. She develops a dual model that incorporates false belief, discussed in chapter five, subsection 2.1.
[82] To this it might be, rightly, objected that the prospective intervener might be wrong, and that they might have the false belief. The correct response to this, however, is not to avoid a basic commitment to truth, but to admit that these situations raise complex epistemic questions. Chapter six, on Meno's paradox, grapples with some of these issues. Chapter seven, on humility, indicates why a commitment to objectivity need not result in uncontrolled paternalism: prospective interveners must also acknowledge the truth of their own limits.

the river should be an overriding consideration, if the collapse of the bridge would stop them from reaching it anyway.

It is also possible to reduce cases of coercion to false belief, although this is less obvious. Chapter three showed that in the 'insufficiency' cases coercion or undue pressure can lead to a person not being treated as so 'autonomous' (in the legal sense) that their expressed wishes must be respected. Subsection 2.4 of this chapter showed that these cases are not reducible to dual model autonomy. They can, however, be incorporated into this four-stage model of prospective harm, false belief, necessity, and proportionality. In some cases, especially those in which the person has been deceived, false belief is obviously relevant. For example, perhaps someone is attempting to cross an unsafe bridge because some other person lied to them, saying that the bridge had been fixed. Then, wrestling them to the ground might seem justifiable, and that justifiability clearly rests on the false belief that the third party caused them to have. In other cases, the role of false belief is more subtle.[83] For instance, imagine someone about to cross the collapsing bridge, knowing that it is collapsing, because another person has threatened to kill their family unless they do so. In this case, too, wrestling them to the ground may be justifiable; but examining the example more closely suggests that this justifiability is tied to false belief. Imagine, as part of some police operation, the coercer has been arrested and the family are now safe. In this situation, wrestling the person to the ground seems not only justifiable, but required. The person is, however, operating under a false belief: that their family will die unless they cross the bridge. If, in contrast, nothing is done to safeguard the family, and they really will die unless the person crosses the bridge, then it is not clear that a third party would be justified in preventing them from choosing to sacrifice themselves. Without the false belief, there is no obvious justification for actions against the person's expressed wishes.

A detailed examination of the issue is beyond the scope of this thesis, but analysis centred on false belief may help to understand some of the most controversial cases that the courts face: those concerning some people with severe anorexia nervosa. People with anorexia may understand the information relevant to their decision, and be able to use and weigh it against their values, but those values may themselves be symptomatic of the anorexia.[84]

[83] Killmister calls these two kinds of false belief version 1 and version 2 (n 81), 526.

[84] Jacinta OA Tan and others, 'Competence to Make Treatment Decisions in Anorexia Nervosa: Thinking Processes and Values' (2006) 13 Philosophy, Psychiatry and Psychology 267.

This cannot be usefully addressed by dual theories: although people are sometimes alienated from the anorexic values, often they endorse them, or have complex ambivalent relationships to them.[85] Assessing a person's values as part of competence assessments, however, raises fears that the values of some groups in society, for instance Jehovah's Witnesses, may be considered incompetent merely due to their statistically unusual beliefs.[86] Putting the problem into the terms of false belief can clarify the issues here. Killmister gives the example of someone who drinks cyanide, knowing it is cyanide and that it will kill them, because they believe that they do not deserve to live after shaming their family with an extramarital affair.[87] Here the person's false belief is about their own moral worth. This suggests an analogy with the anorexia cases. If acting against an anorexic person's expressed wishes is justifiable when they understand the relevant information and can weigh it against their values, then it is because they falsely believe thinness to have more value than it actually does and fatness to have more disvalue than actually it does.

Fears that thinking in terms of false belief about values might necessarily lead to unbridled paternalism would be overblown and self-defeating, yet understandable. Such fears would be overblown because satisfying the false belief criterion alone would not be enough to justify actions against a person's expressed wishes. Any such actions must also avert prospective harm and be necessary and proportionate. As noted in subsection 2.3, judges sometime find someone's beliefs to be false, but nevertheless respect them. In other words, false belief is necessary to justify an action against a person's wishes, but it is not sufficient. Beyond this, basing opposition to paternalism on a claim that there is no truth or falsehood in morality risks self-defeat. If there is no truth in morality, then a claim that 'unbridled paternalism is wrong' cannot be true.[88] All the same, suspicion of acting against a person's wishes because they have a false belief is understandable. Indeed, value-judgments merely present an obvious example of a pervasive problem: knowing what is true is often difficult.

---

[85] Tony Hope and others, 'Anorexia Nervosa and the Language of Authenticity' (2011) 41 Hastings Centre Report 19.

[86] Tan (n 84) 286.

[87] Killmister (n 81) 526-527. Killmister does not believe that such 'tertiary' false beliefs impede autonomy because 'there is no reason to believe that [the person] is mistaken about what she wants to do'. As she notes, however, the person does have 'a false belief about her own moral worth'. Regardless of the, largely semantic, argument about what 'autonomy' is, it is not clear whether she believes such tertiary false beliefs may sometimes form part of the justification for actions against a person's expressed wishes.

[88] In this regard, such an attitude would be an example of what Midgley calls 'dogmatic scepticism', which instead of 'habitually asking questions' is 'so sure that there are no answers' that it 'simply issues denials instead'. Mary Midgley, *Can't We Make Moral Judgements?* (first published 1991, Bloomsbury 2017) 2.

Just because something is difficult, however, does not mean that it can be avoided; and difficulties establishing the truth in particular cases cannot justify pretending to a radical scepticism entirely removed from everyday life.[89] They merely justify proceeding slowly and carefully. For this reason, the difficulty of establishing the truth in a situation is discussed in more detail in chapters six and seven.

False belief and harm are both necessary for actions against a person's expressed wishes, but even the combination of false belief and harm is not sufficient. A coincidence of false belief and prospective harm is not enough to justify actions against the person's expressed wishes. If someone was knowingly about to cross the unsafe bridge, and also happened to believe that snakes are mammals, then their poor grasp of biology would not help to justify wrestling them to the ground. There must be some degree of causation between the false belief and the expressed desire. The person must desire the thing because of the erroneous belief. There are further constraints on actions against a person's expressed wishes: they must also be both necessary and proportionate.

3.3 Necessity

In the bridge example, Mill clearly envisages that the action against the person's expressed wishes must be necessary to avert the prospective danger. He requires that there be 'no time to warn him of his danger'.[90] This requirement is generalizable to all actions against a person's wishes: if it is possible to avert the prospective harm caused by the false belief without contradicting the person's expressed wishes, then that is what should be done. In this respect, acting against a person's expressed wishes should be considered a last resort.

Arguably, this understanding already permeates the current law,[91] if not its implementation.[92] It is not enough that a person fail to understand, retain, use, or weigh some relevant information for them to be found to lack capacity. 'All practicable steps' must

[89] A remark from one of Hume's dialogues is appropriate here: 'Whether your scepticism be as absolute and sincere as you pretend, we shall learn by and by, when the company breaks up: We shall then see, whether you go out at the door or the window; and whether you really doubt, if your body has gravity, or can be injured by its fall'. David Hume, *Dialogues Concerning Natural Religion* (first published 1779, Dorothy Coleman ed, Cambridge University Press 2007) [132], 5.

[90] Mill (n 65) 158.

[91] MCA s1(6): 'Before the act is done, or the decision is made, regard must be had to whether the purpose for which it is needed can be as effectively achieved in a way that is less restrictive of the person's rights and freedom of action'.

[92] Select Committee on the Mental Capacity Act 2005, *Mental Capacity Act 2005: Post-Legislative Scrutiny* (2013–14, HL Paper 139) paras 79–83.

be taken to help them to do those things.[93] In terms of false belief, it is not enough that the person has a false belief about the world. It must be 'impracticable' to persuade them that the belief is false.[94] Something similar occurs under the inherent jurisdiction. In *DL*, Lord Justice McFarlane states:

> …there can be no power of public intervention simply because an adult proposes to make a decision, or to tolerate a state of affairs, which most would consider neither wise nor sensible. There has to be much more than simply that for any intervention to be justified: and *any such intervention will indeed need to be justified as necessary and proportionate*.[95]

Although, as already noted, practice does not always reflect the law in this regard, a necessity requirement is defensible. Any justification for acting against a person's expressed wishes based on prospective harm and false belief collapses if an alternative that respects their wishes would equally avoid the harm. Furthermore, for justification to hold, 'necessity' must be understood in a tightly defined way.[96]

For a justification of the type 'I failed to respect Laura's currently expressed wishes because it was necessary' to work in this context, the word 'necessary' must directly connect to the other parts of the justification: the prospective harm and the false belief. In other words, the proposed action must be necessary to avoid the harm that will be caused by the false belief. It is, however, easy to slip from this to using 'necessity' in a less precise manner. Then purported justifications can take a looser form: 'Laura falsely believes that she will be safe living alone at home, so it is necessary to act against her express wishes'. Justifications of this type take too much for granted in two ways. First, 'safe living at home' is extremely

[93] MCA ss1(3), 3(2)

[94] An incidental side-effect of this four-stage justification concerns the relationship between 'autonomy' in the legal sense and the provision of information. Sufficient information is clearly required to be considered 'autonomous' in law: *Chester v Afshar* [2004] UKHL 41, [2005] 1 AC 134 [24], [57], [77]; *Montgomery v Lanarkshire Health Board* [2015] UKSC 11, [2015] AC 1430 [68], [116]. At the same time, however, there is no suggestion in domestic law that it is appropriate to override a person's expressed wishes simply because they lack relevant information (indeed ss1(2),1(3), and 2(1) of the MCA would seem to forbid it). Why not, if they are 'not autonomous'? This justification makes things clearer. A person who has not been provided with relevant information may have false belief, but actions against their expressed wishes are nevertheless not necessary if they can be given that information.

[95] *DL* (n 60) [76] (emphasis added).

[96] The idea of a 'restrictive' version of the doctrine of necessity is given as a 'cursory suggestion' in Piers Gooding and Eilionóir Flynn, 'Querying the Call to Introduce Mental Capacity Testing to Mental Health Law: Does the Doctrine of Necessity Provide an Alternative?' (2015) 4 Laws 245.

broad. Usually, a person will be safe living at home in some conditions and not in others. For the statement to be valid, a silent external premise must be made explicit and tested: 'Laura believes that she will be safe living alone at home, *but we cannot practicably make her home safe for her*'. In other words, it is necessary to test whether the world can be changed to fit the person's beliefs. If it can, then they are not necessarily false. Second, and similarly, there is a silent internal clause. The justification is only valid if Laura's false belief cannot be changed. Once again, this silent premise must be made explicit and tested: 'Laura falsely believes that she will be safe living at home, *and she cannot be persuaded otherwise*'. If she can be persuaded to change the false belief that her desire is based on, and consequently the desire, then acting against her wishes is not necessary.[97] The common law doctrine of necessity has been criticised as being too 'broad',[98] but that is not exactly right. It is narrow enough, but applied with insufficient rigour. Even the classic statement of the doctrine in *Re F (Mental Patient: Sterilisation)* requires that the person 'cannot by reason of mental disability understand the nature or purpose of an operation or other treatment' that is carried out 'in order either to save their lives, or to ensure improvement or prevent deterioration in their physical or mental health'.[99] Allowing for the contextual differences, these are the conditions of false belief and prospective (relative) harm. In that case, and since, however, there has been a repeated tendency to take both the social context and a person's 'inability to understand' as fixed givens, immune to change.[100] This is not a problem with 'necessity' that renders it unresponsive to context. It a failure to take necessity seriously enough, a failure so severe that the entire justification for acting against the person's expressed wishes collapses.

3.4 Proportionality

A final criterion is implicit in Mill's bridge example. For example, imagine Alice is about to cross a toll bridge, but Jenny, observing, knows that a free bridge is only a half a mile downriver. Once again, there is prospective harm, the cost of the toll, and a false belief, that this is the best crossing. If, furthermore, Alice is walking fast, then tackling her to the

[97] As noted above (n 80), in cases of an inability to 'use and weigh' the relevant false belief is a belief about the relative salience of various factors at play in a situation.
[98] Gooding and Flynn (n 96) 259.
[99] [1991] UKHL 1, [1990] 2 AC 1, 55.
[100] K Keywood, '"I'd Rather Keep Him Chaste." Retelling the Story of Sterilisation, Learning Disability and (Non)Sexed Embodiment' (2001) 9 Feminist Legal Studies 185.

ground may be necessary to stop her going through the toll booth. Nevertheless, unlike in Mill's original example, wrestling her to the ground does not seem justifiable. It seems over the top, disproportionate. This suggests that, in addition to being necessary to avoid the prospective harm, the action against the person's expressed wishes must also fit the overall circumstances in a broader sense. This last sentence is deliberately vague. For lawyers, a proportionality clause is easily taken for granted as a kind of professional common sense, but of the criteria discussed here it may be the most complex.

The most obvious of sort of proportionality is that the action should be proportionate to the envisaged harm that it avoids. In the bridge example, grabbing the person seems more defensible to save their life than to save them from a small toll. Similarly, overriding a person's refusal of medical treatment is more likely to be justifiable if it is necessary to stop them from dying than if it will merely stop them from suffering occasional stomach cramps. This is not, however, the only sense in which proportionality is relevant to the justification of actions against a person's expressed wishes. Consider, for instance, a young woman who vehemently does not wish to have children, but cannot be supported to understand that sex can result in pregnancy, or persuaded to use contraception. It is possible that the conditions of prospective harm, false belief (that she will not risk pregnancy by having unprotected sex), and necessity might be met in such a case. This does not mean that all actions against her expressed refusal of contraception would be proportionate. Usually, a contraceptive patch, because it reversible and less invasive, would be more proportionate than sterilisation.[101] Yet the prospective harm avoided in both cases is identical. The proportionality in question is not just that of the act to the prospective harm but also of the act to the person's life. Sometimes, that might be a test which no available action against the person's expressed wishes will satisfy.

It is at this point, the proportionality of an act to the person's life, that dual theories of autonomy become relevant. Mill is right to insist that humans have extremely diverse opinions and 'modes of life' and that this diversity is a good thing;[102] and dual theorists like Christman are similarly right to insist that even if we are 'muddling through…this difficult

[101] *The Hospital Trust v Miss V* [2017] EWCOP 20 [11] shows that sterilisation is still sometimes suggested even when it would clearly be disproportionate. In 'rare' cases sterilisation may still be found more proportionate than other forms of contraception in the current system, due to the particular risks that a particular woman faces: *The Mental Health Trust v DD (no 4)* [2015] EWCOP 4 [11]-[12] (Cobb J).
[102] Mill (n 65) 121-122.

life', 'it is ours' to muddle through.[103] What is proportionate in one person's life may not be proportionate in another person's; and sometimes a person's reflective endorsement, alienation, or even ambivalence will be relevant to whether or not a particular action against their expressed wishes will be proportionate. To return once more to *Wye Valley*, amputating Mr B's leg against his wishes would have been disproportionate. Those wishes were clear, consistent for over a year, and linked to patterns of belief that he had lived by for decades.[104] Amputating the leg of another man with gangrene who was also refusing treatment because they did not 'understand the reality of his injury' could,[105] however, be proportionate: if, for instance, that man profoundly wished to live, instead of showing 'fortitude in the face of death'.[106] A person's reflective attitude to their own relevant desires can be relevant when justifying actions against their expressed wishes. In this, dual theorists are correct. It is not, however, of any central importance. Unless the person has a relevant false belief that might prospectively cause them harm, and neither the world nor their belief can be changed to remove that risk, then the person's reflective attitude is simply irrelevant. An action against their expressed wishes would be unjustified anyway. Furthermore, even in the final proportionality stage, dual model autonomy is only one factor among many. In Mr B's case, its presence stopped an action against his expressed wishes. It does not follow from this that actions against a person's expressed wishes are always proportionate if this sort of reflection is absent, or are always disproportionate if this sort of reflection is present. As seen at the start of this subsection, proportionality also includes an evaluation of the proposed act and its consequences. Furthermore, even the evaluation of the person as part of the proportionality requirement is unlikely to be reducible to dual model autonomy alone. Chapter three showed that in best interests assessments judges also consider subjectivity and intelligibility: the impact on the person of disregarding their wishes, and the degree to which they are rational.[107] It is hard to see why they are wrong to do so, but these things are not part of dual model autonomy. In other words, dual model autonomy is part of proportionality, but so are other things.

[103] Christman (n 2) 245-246.
[104] *Wye Valley* (n 28) [19], [37], [43].
[105] *Ibid* [34(3)].
[106] *Ibid* [44].
[107] Chapter three, subsection 4.2.

3.5 The difficulty of this justification

This section suggests a four-stage justification for actions against a person's expressed wishes: prospective harm, false belief, necessity, and proportionality. This justification has one important characteristic that reappears in several forms. In almost every respect, it makes things more difficult. This is fitting for a process that began, in the introduction to chapter two, affirming the comment by Kierkegaard, writing as 'Climacus', that 'the crux is exactly the relationship of the subject',[108] not in the development of theoretical systems. Kierkegaard, or at least Climacus, wished 'to make it difficult, as difficult as possible, *yet without making it more difficult than it is*'.[109] That is the point here. Actions against a person's expressed wishes are extremely common, so common that they often become normal. General Comment 1 by the UN Committee on the Rights of Persons with Disabilities rightly challenges that complacency,[110] but even those that wholeheartedly endorse the Comment sometimes treat the job of moral justification as easy. For example, they suggest that a third party can easily determine which 'verbal expressions… represent the true will and preferences of an individual'.[111] This justificatory complacency is indefensible. When even the carefully developed dual models of philosophers like Christman cannot carry this weight, then throwaway comments about 'true will' or 'best desires' amount to nothing more than platitudes. It is better to leave the difficulties in plain sight.

One simple but fundamental difficulty is that nothing has been proven here. In essence, this chapter and the last have reflected on the existing practice of the courts to produce an idealised account of what should be done. Nothing has been derived from unquestionable moral absolutes. Probably, nothing can. As McDowell says, the idea of 'suitable first-natural facts' that can 'force themselves on us in a way that would bypass the need for thought' is a 'fantasy'.[112] There is little choice but to start from where we are, in the middle of things, admitting not only that we may be wrong but also that the tools we use for discovering wrongness may be wrong. In other words, rather than aiming for the elegant simplicity of

[108] Søren Kierkegaard, *Concluding Unscientific Postscript to the Philosophical Crumbs* (first published 1846, Alastair Hannay tr, Cambridge University Press 2009) 33.
[109] *Ibid* 320 (emphasis added).
[110] UN Committee on the Rights of Persons with Disabilities, 'General Comment No 1' (2014) UN Doc CRPD/C/GC/1.
[111] Eilionóir Flynn and Anna Arstein-Kerslake, 'Legislating Personhood: Realising the Right to Support in Exercising Legal Capacity' (2014) 10 International Journal of Law in Context 81, 98.
[112] John McDowell, 'Two Sorts of Naturalism' in Mind, Value, and Reality (Harvard University Press 1998) 189.

theory, this section seeks only 'as much understanding as it can'.[113] As Williams says, reflection will always take things for granted, but 'it must know it will do that'.[114]

The justification also faces practical difficulties. In particular, the question of disability neutrality is pressing. The UN Convention on the Rights of Persons with Disabilities ('the UNCRPD') demands state parties 'recognize that persons with disabilities enjoy legal capacity on an equal basis with others in all aspects of life'.[115] The requirement in the MCA that incapacity be 'because of an impairment of, or a disturbance in the functioning of, the mind or brain',[116] is almost certainly not compliant with this part of the Convention; and, indeed, it seems hard to justify.[117] In contrast, the four-stage justification outlined here does not formally breach the Convention. It does not distinguish between false beliefs that are caused by disabilities and those that are not. All the same, it must be suspected that those with mental disabilities are more likely to be found to have false beliefs than others, and discrimination in the Convention includes distinctions which have the '*effect* of impairing or nullifying the recognition, enjoyment or exercise, on an equal basis with others, of all human rights'.[118] This is an important problem, which should be responded to in a combination of three ways. The first response is to accept it. Nothing in the four-stage justification guarantees disability neutrality in practice, and it should not be suggested that it does. Disability discrimination is a complex, deeply rooted, and constantly shifting socio-cultural phenomenon; and to pretend that it can be removed just by developing the right legal constructs would be absurd. The second response is to push back against the idea that mental disability necessarily entails false beliefs. For instance, 'depressive realism' is now a well-studied phenomenon: people with mild or moderate depression generally have more accurate self-assessments than control groups with regard to their 'relationships, state of health, knowledge and control over external events'.[119] In other words, if it is too easily assumed that those with mental disabilities have false beliefs, then the way to correct that

[113] Bernard Williams, *Ethics and the Limits of Philosophy* (first published 1985, Routledge 2011) 130.
[114] *Ibid.*
[115] Convention on the Rights of Persons with Disabilities (adopted 13 December 2006, entered into force 3 May 2008) 2515 UNTS 3 Art 12(2).
[116] MCA s2(1).
[117] Peter Bartlett, 'The United Nations Convention on the Rights of Persons with Disabilities and Mental Health Law' (2012) 75 Modern Law Review 752.
[118] UNCRPD Art 2.
[119] Lisa Bortolotti and Magdalena Antrobus, 'Costs and Benefits of Realism and Optimism' (2015) 28 Current Opinion in Psychiatry 194, 197.

is to challenge the assumption, not to pretend that false beliefs are always irrelevant to the law. The final response to the question of effective disability-neutrality depends on the first two. If a society is not otherwise discriminatory against people with disabilities, and there is no assumption that those with mental disabilities necessarily have false beliefs, but a higher proportion of people with disabilities than without have decisions made against their express wishes, then that may be acceptable. The four-stage justification attempts to outline when it is right to act against someone's expressed wishes, regardless of disabilities. If, and only if, all sources of discrimination are eliminated and all of the decisions are justified on these grounds, then a higher incidence of decisions against the expressed wishes of those with mental disabilities may be defensible. In terms of the UNCRPD, it may be necessary to 'achieve de facto equality of persons with disabilities'.[120]

Another broad area of difficulty emerges from the four-stage justification. All that it does is justify actions against a person's expressed wishes. It says nothing about how to determine what a person's wishes are. These are two things that the MCA, to some extent, brings together. Unless a person can understand, retain, use, and weigh information then they are not able 'to make a decision for himself'.[121] To an even greater extent, as discussed in chapter two, dual models differentiate between a person's deep and shallow desires to determine what the person 'authentically' wants. Interpretation is part of these models. Furthermore, there is another large area that this justification does not touch. This chapter, and indeed thesis, does not discuss justification when a person cannot express a desire. The MCA simply treats this as another form of incapacity;[122] and some responses to the UNCRPD have also treated it as a continuum.[123] This is perhaps unsurprising. In practice, determining whether someone has wishes, interpreting those wishes, and determining whether those wishes should be respected all run alongside one another. They are simultaneous processes. Nevertheless, they are not the same; and, more importantly, the difficulties with each are not the same.

Any justification based on false belief also creates interpersonal difficulties. It effectively requires saying to the person whose wishes are been overridden not only 'we will not do

[120] UNCRPD Art 5(4).
[121] MCA s3(1).
[122] MCA s3(1)(d).
[123] Michael Bach and Lana Kerzner, *A New Paradigm for Protecting Autonomy and the Right to Legal Capacity* (Law Commission of Ontario 2010) 98-99.

as you say' but 'we will not do as you say because you are wrong'. As with the other forms of difficulty this is a good thing. First, it should not be comfortable for anyone to make decisions against the expressed wishes of another person. If it is too easy, then it will be done without sufficient care. Second, and relatedly, one reason it is uncomfortable to tell another person that they are wrong is because it often hard to be sure about the matter. Sometimes it is not the individual whose expressed wishes are in question who has a false belief, but the individual who proposes to override those wishes. This point is further explored in chapters six and seven.

The final difficulty is that this justification is not neutral between different conceptions of the good. This is most striking with false belief. It demands that the person has a basic orientation towards the truth. Not everyone always does. As Frankfurt observes, some people, at least sometimes, do 'not care whether the things [they say] describe reality correctly'.[124] Demanding some orientation to truth, then, even if it is value-neutral with regard to the contents of different beliefs, is not value-neutral with regard to the process by which a person reaches those beliefs.[125] Similarly, 'necessity' and 'proportionality' may seem neutral in the abstract, but in application they will almost certainly be value-laden. With necessity, whether or not the alternatives to acting against a person's expressed wishes are socially practicable will almost certainly rely on complex mixtures of fact and value. For instance, if a person could safely live at home with eight visits by carers a day, but the local authority will only fund three, then it may be necessary to move them into residential care against their expressed wishes. The local authority's funding decision will ultimately be based on its prioritisation of different levels and types of need, and that prioritisation, a judgement about what is important, will necessarily be value-laden. Indeed the larger decision about what funding the local authority will have ultimately reflects the values of Parliament and the electorate, not of the person in need of care. The role of values is even more obvious at the proportionality stage. How important should the various factors being considered in the decision be? The answer cannot be to use the person's own values alone as a measure. That route either fails to justify ever acting against a person's honestly expressed wishes or assumes that the person's current desire does not reflect some deep,

[124] Harry G Frankfurt, *On Bullshit* (first published 1986, Princeton University Press 2005) 56.
[125] Laura Davy, 'Philosophically Inclusive Design: Intellectual Disability and the Limits of Individual Autonomy in Moral and Political Theory' (2015) 20 Hypatia 132, 139.

consistent self. The latter option is another dual model, and faces all of the problems discussed in chapter two. In conclusion, no part of this four-stage justification is value-neutral. That, however, is good. As Ruth Anna Putnam says, 'nonmoral facts and moral facts are so intimately interwoven' that, despite its popularity in some academic circles, the fact/value distinction can 'bear hardly any philosophical weight at all'.[126] To apply the general point, overriding another person's wishes, or choosing not to, is unavoidably morally difficult, and it does no good to pretend otherwise. It is better to be honest, and that is fitting. The most striking value commitment of this justification is to truth itself. Why is truth valuable? Put simply, it orientates us to reality; and, as Taylor says, 'what is real is what you have to deal with, what won't go away just because it doesn't fit with your prejudices'.[127]

4. Conclusion

This chapter has two parts. The first compares dual theories to the current practice of the Court of Protection, and finds that, although something like a dual model can sometimes feature in best interests decisions, the bulk of the court's concern is with the person's relationship to the world, not with their relationship to their own desires. The second part takes this fundamental orientation of the court as defensible, and uses Mill's bridge example to develop a justification for actions against a person's expressed wishes that requires prospective harm, false belief, necessity, and proportionality. That account makes nothing easy and many things, such as the role of values, more difficult; but not, in Kierkegaard's words, 'more difficult than it is'.[128]

The four-stage justification presents an alternative to value-neutral dual-model autonomy. It is not, however, the only alternative. The next chapter contrasts it to two other types of models of autonomy, which have both appeared under the banner of 'relational autonomy'; and explores the implications for the ethics of mental capacity law.

[126] Ruth Anna Putnam, 'Weaving Seamless Webs' in David MacArthur (ed), *Pragmatism as a Way of Life: The Lasting Legacy of William James and John Dewey* (Belknap 2017) 71.
[127] Charles Taylor, *Sources of the Self: The Making of the Modern Identity* (Harvard University Press 1989) 59.
[128] Kierkegaard (n 108) 320.

Chapter Five

'The elegant, monistic pattern is simply unusable'[1]

1. Introduction

This chapter, like the last, is in two main sections. Section two examines whether alternative models of autonomy, under the banner of 'relational autonomy', can succeed in plausibly distinguishing between when it is and is not permissible to act against a person's expressed wishes. Any that can do so would be a viable alternative to the four-stage justification in the last chapter that requires prospective harm, false belief, necessity, and proportionality for such actions to be legitimate. The section, however, concludes that relational models are no better fitted to this application than the dual models discussed in chapter two. This conclusion allows a broader one to be tentatively drawn. If neither classic dual models nor their most prominent competitors can usefully discriminate between when it is and is not permissible to act against a person's expressed wishes, then it seems likely that no 'model of autonomy' can do so. Section three endorses this conclusion, and takes it further. Models of autonomy are models of the person; or, in the case of one type of relational theory, models of society. There is, however, more to the world than the person, and more to the world even than society. There is more to the situations that mental capacity law operates within, too. Even the simplified example in the last chapter required an unsafe bridge, a real thing that was neither person nor society. It is unlikely that any unitary concept, like autonomy, could ever adequately capture the complexity of persons; but it would be absurd to think that one could ever adequately capture the complexity of the world. As Midgley says, 'the elegant, monistic pattern is simply unusable'.[2] In other words, mental capacity law requires a mode of ethics that orientates itself towards lived complexity, not towards abstract simplicity. In this respect, section three argues that there is still much to be learnt from a group of female philosophers who studied at Oxford during the Second World War.

[1] Mary Midgley, *The Ethical Primate: Humans, Freedom and Morality* (Routledge 1994) 90.
[2] *Ibid.*

2. Relational models of autonomy

Chapter two concluded that no value-neutral dual model of autonomy can plausibly discriminate between when it is and is not permissible to act against a person's expressed wishes. Chapter four concluded that, for the most part, such dual models do not even have the same subject matter as mental capacity law. Accepting these conclusions, however, does not necessarily mean abandoning the idea that a single concept, autonomy, should determine when it legitimate to act against a person's expressed wishes. Perhaps better versions of autonomy are available. This section addresses that possibility. It shows that, as regards the justification of actions against a person's expressed wishes, prominent alternative models of autonomy fare no better than standard dual models. The alternative models are all examples of 'relational autonomy': they all draw attention to the social elements of personhood. They come in two distinctive types. 'Substantive dual models' abandon the commitment to value-neutrality, but keep the dual model structure. They assess a person's process of reflection, but also assess their desires and other characteristics, the *substance* of what they reflect on. 'Pure social views', in contrast, abandon the dual model structure. This section finds that there is no reason to believe either type of theory can plausibly discriminate between when it is and is not permissible to act against a person's expressed wishes in their best interests. Consequently, there is little reason to think that any concept of autonomy can determine these questions.

Applying relational theories in this context requires some reading between the lines. Unlike Christman,[3] those developing these models seldom make the explicit claim that actions against a person's expressed wishes are legitimate if, and only if, they are not autonomous. The author's main interest is usually elsewhere. To return to a distinction discussed in the introduction to chapter three, they generally emphasise appeals to the value of autonomy, arguments of the type 'self-government is good, so we should help Andy be self-governing'.[4] For instance, those applying models of relational autonomy to mental capacity law have made striking arguments that autonomy is valuable, so it should be socially

[3] John Christman, *The Politics of Persons: Individual Autonomy and Socio-Historical Selves* (Cambridge University Press 2009) 162.
[4] George Sher, *Beyond Neutrality: Perfectionism and Politics* (Cambridge University Press 1997) 16.

supported.[5] These arguments, however, have no direct bearing on the issue, central here, of when it is permissible to promote good things, like autonomy itself, for someone against their expressed wishes. If a relational model of autonomy can settle that question, then it must be via an appeal to respect for autonomy, one of the type 'Andy is capable of self-government, and self-government is good, so we should refrain from interfering in his life against his wishes'.[6] In other words, if a relational model is to be able to determine when actions against a person's expressed wishes might be justified it must be capable of producing plausible statements of the type 'when (relational) conditions *A, B, C*… are met, then *P*'s expressed wishes must be respected, but otherwise they need not be'. This section examines the possibility that relational models can be used in this way, without necessarily imputing such an argument to the authors discussed.

2.1 Substantive dual models of autonomy

Substantive dual models add normative criteria to the dual model structure. This can make them appealing. Chapter two found that dual models of autonomy cannot both remain neutral between different conceptions of the good and plausibly justify actions against a person's expressed wishes. It also observed that, in any case, Christman's theory relied on external standards (under the headings of 'constriction', 'pathology', and 'manipulation') that undermined his claims of value-neutrality.[7] In these circumstances, it may seem that dual model autonomy can be rescued by abandoning claims of value-neutrality and admitting that, as well as evaluating the process of reflection, a model evaluates the desire or other characteristic that the person reflects on. Models of this type are 'substantive'. In addition to assessing the process of reflection, they directly or indirectly assess the substance of the person's desires.

Substantive models have been applied to mental capacity law,[8] but care is needed with this literature. It generally assumes a 'gatekeeper' view of the relationship between autonomy

[5] For instance: Beverly Clough, 'What About Us? A Case for Legal Recognition of Interdependence in Informal Care Relationships' (2014) 36 Journal of Social Welfare and Family Law 129. Laura Pritchard-Jones, 'Ageism and Autonomy in Health Care: Explorations through a Relational Lens' (2017) 25 Health Care Analysis 72; Camillia Kong, *Mental Capacity in Relationship* (Cambridge University Press 2017) 5-6 (and throughout, although Kong's core argument is that *capacity,* as well as autonomy, is a relational concept).

[6] Sher (n 4) 16.

[7] Chapter two, subsection 5.3.

[8] Jillian Craigie, 'Capacity, Value Neutrality and the Ability to Consider the Future' (2013) 9 International Journal of Law in Context 4; Catriona Mackenzie and Wendy Rogers, 'Autonomy, Vulnerability and Capacity: A Philosophical Appraisal of the Mental Capacity Act' (2013) 9 International Journal of Law in Context 37;

(in the legal sense) and capacity, which chapter three showed to be over-simplistic. It also often assumes that capacity assessments assess some form of dual model autonomy, which chapter four showed to be erroneous.[9] For these reasons, this subsection draws directly on the philosophical literature, rather than on existing applications of that literature. It finds that substantive models cannot plausibly distinguish between when it is and is not permissible to act against a person's expressed wishes. These models fail to avoid a problem discussed in chapter two: there is no plausible way of assessing the person's reflection that can legitimate actions against their expressed wishes.[10]

Substantive dual models are not homogenous: there are both 'strong' and 'weak' substantive accounts.[11] A strong substantive account is any dual model that directly imposes normative restrictions on what an autonomous agent can desire or believe. For instance, a strong substantive account may hold that a woman cannot be autonomous if she believes that she must always obey her husband, even if she reflectively endorses that belief. The content of the belief, by itself, is enough to render her non-autonomous. In contrast, weak substantive accounts only indirectly constrain what an autonomous agent can desire or believe. For instance, a weak substantive account may say that an agent is only autonomous if they trust their own abilities when reflecting. On this account, a woman can autonomously believe that she must always obey her husband, but only if she trusts her own ability to decide that question. The normative constraint is on the process of reflection, not directly on the content of the person's belief. Due to the sheer variety of substantive dual models, this subsection examines three: a strong one, a relatively simple weak one, and a more complex weak one.

Stoljar provides an example of a strong substantive account.[12] She gives a dual model, not very different to Christman's, and argues that it fails to align with the 'feminist intuition' that preferences influenced by oppressive norms are not autonomous even if they are

---

Fabian Freyenhagen and Tom O'Shea, 'Hidden Substance: Mental Disorder as a Challenge to Normatively Neutral Accounts of Autonomy' (2013) 9 International Journal of Law in Context 53; Kong (n 5).

[9] Chapter four, subsections 2.1, 2.2.

[10] Chapter two, section 6.

[11] Natalie Stoljar, 'Autonomy and the Feminist Intuition' in Catriona Mackenzie and Natalie Stoljar (eds), *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self* (Oxford University Press 2000) 94, 107. The terminology in this literature is far from consistent. For instance, Kong makes a similar distinction between 'procedural-relational' and 'substantive-relational' theories (n 5, 58).

[12] *Ibid.* Other examples are Joseph Raz. *The Morality of Freedom* (Oxford University Press 1986) 378-390; Marina Oshana, 'How Much Should We Value Autonomy?' (2003) 20 Social Philosophy and Policy 99.

reflectively endorsed.[13] For this reason, she adds the requirement that an autonomous agent cannot have internalised 'false norms', such as the belief that 'women should not actively desire sex'.[14] The idea of false norms makes this model superficially similar to last chapter's four-stage justification, which includes false belief. Both draw attention beyond the relationship between the person and their own desires, to the relationship between the person and the world.[15] There are, however, two major differences between Stoljar's theory and the four-stage justification, each of which highlights a problem that strong substantive theories have, if used to determine when it is acceptable to act against a person's expressed wishes.

The first difference between Stoljar's model and the four-stage justification is that the scope of 'false norms' is narrower than 'false belief'. The last chapter drew on Killmister's three-part classification of false belief: a person can be mistaken about the act that they are performing; about what follows from the act; or about the norms associated with the act.[16] Stoljar's 'false norms' are identical with the last of these categories, and this immediately raises practical issues. Although understanding the nature of a particular act and what will follow from it can be difficult, it is usually easier than discovering incontestably true norms. Indeed, the norms Stoljar points out as 'false', such as 'pregnancy is an expression of "real" womanhood',[17] have been considered true by the vast majority of women throughout history, and are still believed by many today. Kong fears that strong substantive dual models collapse into the claim that only those with 'the same sociocultural values' as the speaker are autonomous;[18] and this fear seems well grounded, as a closer look at what Stoljar's theory would mean in practice reveals.

Stoljar's model cannot justify actions against someone's expressed wishes without collapsing into circularity. She bases her claim about autonomy on the intuition that certain

[13] (n 11) 95.
[14] *Ibid* 109.
[15] Indeed, this emphasis has led Benson to comment that this part of Stoljar's theory is not about autonomy, whether or not the person is self-governing, but 'orthonomy', whether or not the person is right. Paul Benson, 'Feminist Intuitions and the Normative Substance of Autonomy' in James Stacey Taylor (ed), *Personal Autonomy: New Essays on Personal Autonomy and its Role in Contemporary Moral Philosophy* (Cambridge University Press 2005) 124, 132.
[16] Suzy Killmister, 'Autonomy and False Beliefs' (2012) 164 Philosophical Studies 513, 526-7.
[17] Stoljar (n 11) 109.
[18] Kong (n 5) 72-74.

norms are not compatible with it.[19] This sort of working from intuitions is common among autonomy theorists.[20] It gives an argument of the following form: 'people like *X* are not autonomous, so autonomy is *not-X*'. Applying a theory, however, works in the opposite direction. It takes the following form: 'autonomy is *not-X*; so this person, who is *X*, is not autonomous'. The combination is catastrophic. For instance, imagine Nina wishes to remove Billie from an abusive relationship against Billie's wishes. She believes that she can do this because Billie is not autonomous. She believes that Billie is not autonomous because Billie has internalised the norm 'I must obey my husband, no matter what'. Why does Nina believe this norm impairs Billie's autonomy? Because people like Billie believe it, and they are not autonomous. Why are they not autonomous? Because they have internalised the norm 'I must obey my husband, no matter what'. In application, Stoljar's theory has absolutely no meaningful content, so there is nothing that Billie could meaningfully contest. It is worth contrasting this with the four stage justification. In it, Nina must say to Billie, 'you are wrong in a way that will harm you, and acting against your wishes is necessary and proportionate'. Each of the four stages – harm, false belief, necessity, and proportionality – has content that Billie can contest.

Strong substantive theories have a second, more fundamental, difference from the four-stage justification; and it, too, reveals problems that they will face in application. In them, to be autonomous, someone must still endorse their desire on hypothetical reflection, even if they are not guided by false norms. The four-stage justification has no such requirement. If the person does not have a relevant false belief, then no action against their expressed wishes is ever permissible.[21] As a consequence, Stoljar's model, but not the four-stage justification, requires some method of assessing reflection that can legitimate actions against a person's expressed wishes in the absence of false belief. There is no such method. This, however, is a problem that her model shares with weak substantive dual theories, so it is examined after introducing them.

Benson's influential weak substantive theory, which applies only indirect constraints on what an autonomous person can believe or desire, is a good introduction to the type. In it,

[19] Stoljar (n 11) 95, 108-9.
[20] See, for example, Freyenhagen and O'Shea (n 8) 61-64, Kong (n 5) 66, 166.
[21] If they *do* have a false belief, then real reflection may render acts against their expressed wishes illegitimate anyway, but that is another issue. See chapter four, subsections 2.3 and 3.4.

a person must have 'a certain sense of their own worthiness to act' in addition to dual model reflective endorsement.[22] This avoids the circularity of Stoljar's account. It gives a reason why a norm like 'I must obey my husband' might undermine a person's agency, and some idea of the conditions in which it will do so. For instance, in this theory, Billie would only lack autonomy if her belief in the norm of husband-obeying stopped her from considering herself worthy to make a valid decision about whether to stay or go. If she trusts herself to make the decision, and endorses the decision on reflection, then even a decision to stay in the abusive relationship is autonomous. Benson's work draws attention to the extent to which abusive relationships can undermine a person's trust in themselves, and this is valuable. It does not, however, plausibly distinguish between when it is and is not permissible to act against a person's expressed wishes. In application, the self-trust part of his model collapses into a 'bare reflection' standard,[23] one that any expressed wish will meet. If, for instance, Billie says 'I am staying in the relationship, because I must obey my husband' she has, simply by expressing a wish, taken herself to be worthy enough to reach a conclusion. As Stoljar points out, there is a difference between general feelings of unworthiness and feeling unworthy to be an agent.[24] If, as in Benson's model, only feeling incapable of being an agent is relevant, then the model cannot justify actions against a person's expressed wishes. By expressing wishes, the person has trusted themselves as an agent. If, in contrast, a model takes other feelings of unworthiness as enough to render someone non-autonomous, then it is a strong substantive theory that directly constrains what an autonomous agent can believe, so it face the same problems as Stoljar's. Beyond this, Benson's theory faces a second problem. In it, even if the person has self-trust they must also reflectively endorse a desire for it to be autonomous; so, like Stoljar, Benson must provide some plausible method of assessing reflection. This he cannot do, but as this problem is also shared by a more plausible weak substantive theory, it will be addressed after introducing that theory.

A more promising weakly substantive model has been proposed by Killmister. For her, autonomy requires reflective competence and endorsement,[25] self-trust in a manner similar

[22] Paul Benson, 'Free Agency and Self-Worth' (1994) 91 The Journal of Philosophy 650.
[23] Chapter two, subsection 5.2.
[24] Stoljar (n 11) 108 – 109.
[25] Killmister (n 16) 519.

to Benson,[26] *and* that the person's relevant beliefs are true.[27] In other words, it is a dual model with two indirect constraints on a person's reflection. They must sufficiently value themselves, and their beliefs must sufficiently accord with reality. As with other substantive dual models, Killmister's faces problems with application. It is a model of the person, not the situation, which means that it omits vast relevant areas. Applied mechanically, it would allow actions against a person's expressed wishes even when complying with their wish would cause no harm, or when the action against their wish is unnecessary or disproportionate. This seems hard to justify. Adding those requirements, however, is effectively to concede a central point of this thesis: autonomy alone cannot determine when actions against a person's expressed wishes are legitimate.

Adding harm, necessity and proportionality requirements to Killmister's model also reveals the central difference between substantive dual models and the four-stage justification; and this, in turn, reveals a problem with all substantive models. Under the four-stage justification, unless the person has a relevant false belief, no action against their expressed wishes is ever legitimate. In contrast, all substantive dual models would also allow actions against the person's wishes if they would not endorse the desire on reflection, even in the absence of false belief. This raises a question also asked in chapter two: how will the person's reflection be assessed?[28] There are no plausible answers that can preserve a justification for acting against someone's expressed wishes in the absence of false belief, as can be shown by returning to the categories from chapter two: real reflection, mere competence, and hypothetical reflection.

If the person is held to a real reflection standard, they must actually reflect. This raises familiar problems.[29] If the model calls for reflective endorsement, most people have not actively endorsed most of their desires on reflection, so most people will be found not to be autonomous most of the time. If, in contrast, it calls for non-alienation on reflection, then a person who has not reflected is not alienated, and so there is no warrant for acting against their expressed wishes. In either case, all that a real reflection standard can justify is prompting actual reflection, but once the person has reflected it supplies no justification for

[26] *Ibid* 527.
[27] *Ibid* 528, she does not think that true beliefs about values, of the type Stoljar requires for autonomy, are necessary, 527.
[28] Chapter two, section 6.
[29] See chapter two, subsection 6.1.

acting against their expressed wishes. If someone prompts reflection by asking Adrian if he should really be returning to his abusive ex-wife, then none of his possible answers justify overriding his expressed wishes. If he says 'she's not abusive, she's just dominant' then he has reflectively endorsed the desire. If it is objected that he is wrong, and Helen really is abusive, then the actual work is being done by the 'false belief' part of the theory. The elaborate dual model structure adds nothing. If Adrian decides not to return to the relationship, reflection has changed his desire, so no action against his wishes is in question. If he says, 'I wish I didn't love her so much' but returns anyway, then some reason is needed for treating his words as a second-order reflection on first-order desires without treating his actions as a third-order reflection on his second-order reflection.[30] Where false belief is not an issue, a real reflection standard fails to plausibly justify actions against the person's expressed wishes. When it is an issue, the dual model adds nothing of practical substance.

If the person is held to a mere competence standard,[31] they must simply be able to reflect. Then, however, the same problem is encountered. In Killmister's model, someone must be able to form an intention, and either endorse or be alienated from that intention on reflection. 'On reflection' includes being guided by self-trust and relevant true beliefs. Self-trust, as seen in the discussion of Benson's model, adds nothing of practical consequence. It collapses into a 'bare reflection' standard. If, however, the issue is false belief then, as with real reflection, the dual model structure adds nothing. The 'false belief' criterion is doing all the work. If neither self-trust nor false belief are in issue, then this leaves a mere competence standard combined with 'bare reflection over time', which is exactly the situation discussed in chapter two.[32] Then, as seen in the 'Elena' example of repudiating a future delirious refusal of treatment, the test no longer necessarily tracks dual model autonomy, so dual model autonomy cannot justify its use. As with real reflection, testing mere competence adds nothing when false belief is the issue, but it cannot legitimate actions against a person's expressed wishes when it is not.

The final possibility for assessing a person's reflection is by hypothetical reflection. Here, substantive dual models face the same problems as other dual models. As discussed in

[30] See chapter two, subsection 5.2 and Gary Watson, 'Free agency' in Gary Watson (ed) *Free Will* (Oxford University Press 1982) 96, 108.
[31] See chapter two, subsection 6.2.
[32] *Ibid*.

chapter two, there are two forms of the hypothetical reflection test.[33] The strong form is 'would P be alienated from, and not authentically endorse, the relevant desire in any imagined world in which they reflected?' As seen in chapter two, this can never justify actions against the person's expressed wishes. It is always possible to construct some story in which the person would authentically endorse the desire on reflection, and requiring that reflection must be guided by true belief and self-trust does not change this. On the strong version of the test, any such story, no matter how absurdly implausible, renders actions against a person's expressed wishes illegitimate. The weaker form of the test is 'is a world in which P reflects and is alienated more similar to the real world than a world in which they reflect and are not alienated?' As discussed in chapter two, 'similarity' in this test is not value-neutral, but this is not necessarily a problem for substantive dual models. They do not claim to be value-neutral anyway. Chapter two, however, also pointed out that 'similarity' is radically indeterminate. It has no content that is not contextual. To return to the example from that chapter, if John's daughter Sally says that John would want to be buried with his parents because he always stopped by their grave on Sundays, then she is not comparing a neutral fact to an established standard. She is both drawing attention to a fact about John's behaviour, and claiming that this fact should count for 'similarity'.[34] For the legitimisation of actions against a person's expressed wishes, this is absurd. There is an obvious way in which a world in which the person reflects and endorses their expressed wish is always more similar to the real world than a world in which they reflect and do not endorse that wish. They still end up expressing the same wish. Imagine I am eating chips. You see me, and say that a world in which I do not want chips is more similar to this world, in which I want chips, than a world in which I want chips. Not content with that, you use this claimed similarity to justify taking my chips away from me, despite my protests. In this example, one of us has certainly become mentally unstuck. It is not, however, me. Hypothetical reflection repeats the same issue found with real reflection and mere competence. Insofar as the false belief is relevant, the dual model adds nothing. Insofar as it is irrelevant, the dual model cannot justify actions against a person's wishes.

Substantive dual models disavow value-neutrality, yet still cannot set a plausible way of assessing a person's reflection that legitimates acting against their expressed wishes in the

[33] Chapter two, subsection 6.3.
[34] *Ibid.*

absence of false belief. When the possibilities – real reflection, mere competence, and hypothetical reflection – are canvassed, a pattern repeats itself. Insofar as the problem is a false belief the person has, actions against their expressed wishes might be justified, but the dual model structure does nothing; but insofar as the person does not have a false belief, the model cannot plausibly justify actions against their expressed wishes. In sum, the dual model is irrelevant.[35] This may be because substantive dual models have the wrong subject matter. For all that they claim to be 'relational', they are models of persons, not situations. The question they ask is 'is this person autonomous?' In contrast, the four-stage justification asks 'is acting against these expressed wishes justified in these circumstances?' This wider focus is not, however, unique to it. It is, to a limited extent, shared by pure social views of autonomy.

2.2 Pure social views of autonomy

Instead of disavowing value-neutrality, like substantive dual models, there is another potential way to defend the idea that a concept of autonomy can determine the legitimacy of acting against a person's expressed wishes: abandon the dual model structure. Instead of locating autonomy in the person's relationship to themselves, such 'pure social views' locate it in the person's relationship to society.[36] The last chapter showed that mental capacity assessments examine the person's relationship to the world, and society is part of the world, so pure social views may fit existing practice better than dual models. Furthermore, without the dual model structure, such models will avoid the problems with reflection discussed in chapter two and the last subsection of this chapter. Together, these two points suggests that pure social views might successfully determine when actions against a person's wishes are acceptable. This subsection shows that they do not fulfil this promise, but the reasons why they fail are themselves revealing.

[35] This does not mean that dual models are entirely irrelevant to mental capacity law. When someone does have a relevant false belief then actions against their expressed wishes may anyway be illegitimate in the presence of real reflection. See chapter four, subsections 2.3 and 3.4.

[36] The phrase 'pure social view' originates in Michael Garnett, 'The Autonomous Life: A Pure Social View' (2014) 92 Australasian Journal of Philosophy 143. Garnett's own view, that 'a person is autonomous simply to the extent to which it is difficult for others to subject her to their wills' (143) does not seem a plausible candidate for determining when actions against a person's expressed wishes are permissible. It would simply hold that such actions infringe autonomy.

Compared to dual models, relatively few pure social views have been developed. The most influential is Westlund's.[37] For her, an autonomous person 'holds herself answerable, for her action-guiding commitments, to external critical perspectives'.[38] In other words, autonomy is not a matter of endorsement or non-alienation on reflection, but of being open to justification and reason. This does not require that someone always agree with criticism. To be autonomous, they merely have to engage with other perspectives, and even this is not an absolute demand. They must only hold themselves answerable to legitimate critics, who have some stake in what the person chooses, to a context-specific degree, which allows for different modes of communication.[39] Westlund believes these limits make her account value-neutral.[40] This claim, however, is unsustainable. In her model, what counts as both 'legitimate' and 'context-specific' is socially determined.[41] In other words, it need not reflect the person's own values. If, for instance, the person does not value her relationship with her social worker, or the practice of talking about her actions with that social worker, but she must nevertheless be willing to engage for her wishes to be respected, then this involves the imposition of values that are not her own. If that is so, though, then Westlund does not offer practical value-neutrality. Her model faces a problem analogous to those faced by dual models. As with 'constriction, pathology, or manipulation' in Christman's model,[42] unanalysed norms do all the real work. Simply because Westlund has neglected to fill out any specifications for 'legitimate' 'context-specific' criticism does not mean that the model can work without such specifications. Once the specifications are provided, however, the model is no longer value-neutral.

It would perhaps be possible to abandon claims of value-neutrality and develop a substantive pure social view of autonomy. Pure social views, however, face a more fundamental problem. Giving reasons for action, and so the practice of being held accountable, can be a way of managing problems with a person's agency. That does not mean that problems with agency are reducible to problems with giving reasons. An old

[37] Andrea C Westlund, 'Rethinking Relational Autonomy' (2009) 24 Hypatia 26.
[38] *Ibid* 35. She remains agnostic about whether autonomy has any other requirement, but is noticeably sceptical of dual models. See also Andrea C Westlund, 'Selflessness and Responsibility for Self: Is Deference Compatible with Autonomy?' (2003) 112 The Philosophical Review 483.
[39] Westlund (n 37) 39-40.
[40] *Ibid* 38.
[41] *Ibid* 40.
[42] Christman (n 3) 147. See chapter two, subsection 5.3.

example illustrates the point. Kierkegaard, writing as Climacus, tells the allegedly true story of a patient who runs away from 'an insane asylum'.[43] Terrified of being returned, he aims to 'completely convince everyone, through the objective truth of what you say, that the matter of your sanity is quite in order'.[44] So when he meets a friend, he just repeatedly says 'the earth is round!'[45] He is quickly returned to the asylum. The story is worth close attention. It is almost universally agreed that the world is round, and this is a silent but necessary premise to any particular reason for acting. Unless the world was round there would be no asylum, no man, and no human actions at all. All the same, talking about it is odd. Kierkegaard's example can be developed. Imagine the man had a more contextual understanding of giving reasons. Then, he might say something along the lines of 'to convince everyone that my sanity is in order, I need to give reasonable reasons for my behaviour, grounded in shared assumptions'. He meets his friend. He says 'Hello. I am greeting you in order to affirm our mutual connection and express my happiness at our meeting. I am explaining my greeting in order to reassure you that my sanity is in order. I am becoming subdued because you are looking at me strangely and I wish you to believe that I am sane…' and so on. The man would be returned to the asylum just as quickly. Something is wrong with his behaviour, but it is not that he is not holding himself answerable. It is that he is holding himself too answerable. The underlying issue is simple. Westlund is right that doubts about a person's agency emerge within a matrix of everyday human practices. She is also right that when such doubts do emerge, they are often managed by requesting and giving reasons. That does not mean that either everyday human practices or a person's agency can be reduced to the giving of reasons. Asking someone for reasons for their behaviour, and so Westlund's need for a person to hold themselves answerable, is a common response to a perception that something is unusual about that behaviour. It is not, itself, the source of the perception that something is unusual.

This conflation between perceived problems with agency and the social practices that manage those perceived problems draws out the fundamental problem with pure social views. Humans are very social creatures, but they are not purely social creatures. Sometimes, someone may be right not to hold themselves answerable to socially legitimate

[43] Søren Kierkegaard, *Concluding Unscientific Postscript to the Philosophical Crumbs* (first published 1846, Alastair Hannay tr, Cambridge University Press 2009) 149.
[44] *Ibid* 150.
[45] *Ibid.*

and context-specific demands for reasons. Fairly often, society is wrong about what is legitimate and about what the context demands. Two hundred years ago, anyone in the United States of America who refused to answer the court when summoned for hiding a fugitive slave would be refusing to hold themselves answerable to socially legitimate, context-specific demands.[46] To modern eyes, such a refusal does not look like a failure of autonomy. It looks like an exemplar of it. By drawing attention to practices of justification and interpretation, pure social views get something important right; but by reducing everything to those practices, they get something important wrong. The last chapter provides the material to say what that is. Dual models focus on the person's relationship to themselves, but the more important issue is their relationship to the world. In a similar manner, purely social views focus on the person's relationship to society, but society is only one part of the world. It is an obvious truth that many people used to think that slavery was acceptable but that few now agree. The more important truth, however, is that slavery was wrong then and is wrong now. Even saying this, however, requires a commitment to truths that are not entirely socially contingent. To return to Murdoch's distinction,[47] sincerity, even sincerity shared by an entire group, cannot ground morality. Ultimately, as she says elsewhere, 'the authority of morals is the authority of truth'.[48] Pure social views do not, as they claim, achieve value-neutrality; but, more fundamentally, by reducing autonomy to a purely social phenomena, they render themselves incapable of responding to those parts of the world that are not purely social. As with substantive dual models, they offer no concept of autonomy can plausibly guide practice.

3. Moving beyond 'autonomy talk'

So far, this thesis suggests that actions against a person's expressed wishes cannot be plausibly justified by value-neutral dual theories,[49] and that alternative models of autonomy offer no prospect of doing any better.[50] One response a situation like this might be to propose some new, better, account of autonomy. The thesis, however, also suggests that although judges do not use the word 'autonomy' in a coherent way, a coherent account

[46] An Act Respecting Fugitives from Justice, and Persons Escaping from the Service of their Masters 1 Stat 302 1793, s4 (USA).
[47] Iris Murdoch, 'Against Dryness' (1961) 16 Encounter 16. Introduced at the start of chapter four, section 3.
[48] Iris Murdoch, *The Sovereignty of Good* (first published 1970, Routledge 2014) 88.
[49] Chapter two.
[50] Chapter five, section 2.

of when they will override or respect a person's expressed wishes can be constructed;[51] and that concepts of autonomy are marginal to the justification of this account.[52] One response this, more complex, situation might be to propose some new unitary concept, other than autonomy, that will decide these questions. This section argues that such a response would be a mistake. The fundamental problem is not the concept of autonomy, but the underlying assumption that ethics can be usefully reduced to a handful of tractable principles. As the title of this chapter, taken from Midgley, says 'the elegant, monistic pattern is simply unusable'.[53] Subsection 3.1 makes this point with regard to the four-stage justification proposed in the previous chapter, showing that it is not a model of the person but an open-ended and non-exhaustive model of situations. Subsection 3.2 places this argument in a broader context, arguing that Midgley and her contemporaries have already pioneered a style of ethics that suits the complexities of mental capacity law. Although legal theory currently draws on a remarkably narrow set of ethical concepts, there is no reason that it must do so. The history of ideas is long and diverse. Subsection 3.3 illustrates this point by showing how the following two chapters, six and seven, each draw on old ideas to address mental capacity law from new angles.

3.1 Situating 'autonomy talk'

This subsection clarifies two points. First, the four-stage justification for actions against a person's wishes is not a model of autonomy; and, second, it cannot be reduced to any single concept. As regards the first point, there are both pragmatic and principled reasons for denying the four-stage justification is a model of autonomy.

It could be suggested that in the four-stage justification a person is 'autonomous' unless the following applies: they risk harm because of a false belief, and overriding their expressed wish is both necessary to avoid that harm and proportionate. There is a strong pragmatic reason to avoid this suggestion. 'Autonomy', in philosophy, is already associated with dual models; and, as shown in chapter four,[54] this sort of autonomy is sometimes relevant to best interests decisions. At the same time, chapter three shows that judges already talk about 'autonomy' in a variety of conflicting ways that often have little to do with the philosophical

[51] Chapter three.
[52] Chapter four.
[53] Midgley (n 1) 90.
[54] Chapter four, subsection 2.3.

use of the word. Beyond these considerations, appeals to the value of autonomy, arguments of the type 'autonomy is good for P, so we should promote it', also appear frequently in both law and philosophy, but they are beyond the scope of this thesis. There is no guarantee that 'autonomy' in those formulations always refers to any of the various concepts discussed here. In these circumstances, calling yet another thing 'autonomy' will not help. The word is already, as Midgley says of the word 'moral', 'like a small carpet, designed to fit a room which has been enlarged', 'wrenched this way and that to cover the bare spaces'.[55]

There is also a principled reason to avoid calling the four-stage justification a 'model of autonomy'. Self-government is marginal to it. Harm and false belief are both about the person's relationship with the world, not about their relationship to themselves. Similarly, 'necessity' examines what else could be done in the world to either change circumstances or the person's mind, and 'proportionality', although it may sometimes include something like dual model autonomy, also draws on a wide range of norms and facts about the world. In other words, actions against a person's expressed wishes are not justified by reference to the person alone, only by reference to the person as part of a shared world. This is not 'relational autonomy'. As seen in section two, relational models still primarily assess the person or, in the case of pure social views, society. The four-stage justification, in contrast, assesses the situation. To call this a 'model of autonomy' would be inaccurate. It is not directed to the question 'is this person self-governing?' but to the question 'can this particular third party justifiably fail to respect these particular expressed wishes in these particular circumstances?'

It is because the four-stage justification assesses an entire situation that it cannot be reduced to any single concept. It explicitly draws on the concepts of harm and, through false belief, truth. Both of these are extremely basic concepts, and it seems unlikely that one can be convincingly reduced to the other. Beyond that, 'proportionality' necessarily draws upon a wide variety of other norms. Chapter three has already shown how best interests assessments draw on the diverse ideas of intelligibility, subjectivity, and authenticity when evaluating the proportionality of actions against someone's wishes. There is no obvious reason to think that this is an exhaustive list. Furthermore, other moral concepts are relevant to these situations, as the detailed discussion of humility in chapter seven shows.

[55] Mary Midgley, 'Is "Moral" a Dirty Word?' in *Heart and Mind* (first published 1972, Routledge 2003) 119.

Admittedly, these considerations do not prove that autonomy cannot be replaced by some single, elegant concept that will distinguish between when it is and is not permissible to act against someone's expressed wishes for their own good. They do, however, make that very ideal, conceptual elegance, suspicious. Why would *that* be the important thing? Other intellectual ideals exist, as the following passage from Bernard Williams, who learnt much from the philosophers discussed in the next subsection,[56] indicates:

> Theory typically uses the assumption that we probably have too many ethical ideas, some of which may well turn out to be mere prejudices. Our major problem now is actually that we have not too many but too few, and we need to cherish as many as we can.[57]

The next subsection fleshes out a way of 'cherishing' ethical ideas. First, however, a potential problem with replacing autonomy, in particular, with a multitude of moral ideas must be addressed. It might be objected that the whole point of thinking about autonomy is to ensure that the person is treated with due respect. In other words, it could be said that 'autonomy' is not about the overall moral assessment of a situation, it is about ensuring the individual is respected within that situation. This is a valid concern, but it assumes too much. Dual models have been elaborated by philosophers for half a century, yet as chapter two and the first part of this chapter show, they are still not practically applicable in what might be considered one of the concept's core domains: the question of when it is acceptable to override another person's expressed wishes. Given this, there are reasons to be suspicious of the very idea that any single concept can ensure people are always adequately respected. Human beings are diverse, and their shared situations are diverse. No single idea can operate like a moral skeleton key, always able to unlock the door to interpersonal respect. To be blunt, life is more complicated than that. To be blunter, a failure to acknowledge that life is intractably complicated is, itself, a moral failing. As Nagel points out, insisting on a 'highly unified conception' leads to 'philosophical mistakes'.[58] These include, pertinently, 'false reductions or to the refusal to recognize part of what is real'.[59] Philosophy influences legal theory, as is occurring with dual models,[60] and yesterday's

[56] Bernard Williams, *Ethics and the Limits of Philosophy* (first published 1985, Routledge 2011)263.
[57] *Ibid* 130.
[58] Thomas Nagel, *The View from Nowhere* (Oxford University Press 1986) 3.
[59] *Ibid.*
[60] Chapter two, section 3.

philosophical mistakes can become today's academic legal mistakes. Then, if legal theory influences practice, we risk a situation that Charles Taylor warns of: one where 'we can't see any more the way these rules fit badly our world of enfleshed human beings, we fail to notice the dilemmas they have to sweep under the carpet'.[61] Taylor, like Williams, learnt that other approaches to ethics are possible from a remarkable group of female philosophers in the generation before his own.[62]

3.2 Learning from the 'Golden Age of Female Philosophy'

In 2013, a letter appeared in *The Guardian* under the provocative title of 'The Golden Age of Female Philosophy'.[63] In it, Mary Midgley responded to an article by Jonathan Wolff that asked why philosophy remained 'the most male-dominated discipline in the humanities'; and pointed out that the University of Oxford during the second world war, when most of the men were away, produced an 'unmatched' cohort of female philosophers, including Elizabeth Anscombe, Philippa Foot, Iris Murdoch, and Midgley herself.[64] To this, Midgley responded 'sorry, but the reason was indeed that there were fewer men about then'.[65] She continued by saying that the problem was not 'men as such – men have done good enough philosophy in the past', but a 'particular style of philosophising that results from encouraging a lot of clever young men to compete in winning arguments'.[66] To this, she contrasted a philosophy that was 'more interested in understanding this deeply puzzling world'.[67] The letter sparked wider interest. It led to 'A Golden Manifesto' in *Philosophy Now* magazine,[68] which attempted to summarise some commonalities of the wartime group; and, more recently, to a research project at Durham University that, for the first time, treats the

---

[61] Charles Taylor, *A Secular Age* (Harvard University Press 2007) 742.

[62] GEM Anscombe was his PhD supervisor, but Iris Murdoch's work seems to have been a greater influence. Charles Taylor, 'Iris Murdoch and Moral Philosophy' in *Dilemmas and Connections: Selected Essays* (Harvard University Press 2011) 3; Matthew JM Martinuk, 'A Fundamental Orientation to the Good: Iris Murdoch's Influence on Charles Taylor in Mark Luprecht (ed), *Iris Murdoch Connected: Critical Essays on Her Fiction and Philosophy* (University of Tennessee Press 2014) 183.

[63] Mary Midgley, 'The Golden Age of Female Philosophy' *The Guardian* (London, 28 November 2013) < https://www.theguardian.com/world/2013/nov/28/golden-age-female-philosophy-mary-midgley> accessed 20 November 2017.

[64] Jonathan Wolff, 'How Can We End the Male Domination of Philosophy?' *The Guardian* (London, 26 November 2013) <https://www.theguardian.com/education/2013/nov/26/modern-philosophy-sexism-needs-more-women> accessed 20 November 2017.

[65] Midgley (n 63).

[66] *Ibid.*

[67] *Ibid.*

[68] Mary Midgley, 'A Golden Manifesto' (2016) 116 Philosophy Now 34; Mary Midgley, 'A Golden Manifesto, Part II' (2016) 117 Philosophy Now 20.

women as a 'group of philosophers pursuing a single, multi-faceted programme'.[69] This is not the place for a full account of that programme. As Taylor says of Murdoch, their work 'is much too rich, and we are much too close to it'.[70] Nevertheless, there are good reasons to think that both the group's diagnosis of the problem with a particular 'kind of philosophy' and their responses to that problem have important lessons for mental capacity law.[71]

Their diagnosis remains startlingly relevant. For instance, in 1964, Murdoch criticises a view of human personality that she finds in both moral philosophy and politics:

> Our personal being is the movement of our overtly choosing will. Immense care is taken to picture the will as isolated. It is isolated from belief, from reason, from feeling, and is yet the essential centre of the self. '*I identify myself with my will*'.[72]

Here, a decade before Frankfurt's seminal article, is a critical sketch of the underlying structure of dual model autonomy by someone who finds it 'alien and implausible'.[73] In both content and tone, Murdoch's criticism is very similar to some recent relational autonomy literature. For instance, there is a clear overlap between the quoted passage and Dodds' argument that 'the conception of autonomy used in bioethics is rationalistic, atomistic, and individualistic'.[74] Relational autonomy theorists, however, make no reference to Murdoch and her contemporaries when making versions of arguments that the older women pioneered over fifty years ago. This may be because their predecessors are more radical than they are, to the point that the older critique can also be applied to relational models of autonomy. Relational theorists diagnose a problem with particular conceptions of autonomy, but the golden age diagnosis is of a global problem with an entire style of philosophy. Murdoch, does not, like the theorists discussed in section two of this

---

[69] Rachael Wiseman, 'Anscombe, Foot, Midgley and Murdoch: A Female Philosophical School?' (Annual Women in the History of Philosophy Lecture, The University of Sheffield, 12 May 2017) < http://www.womeninparenthesis.co.uk/anscombe-foot-midgley-and-murdoch-a-female-philosophical-school/> accessed 20 November 2017.

[70] Taylor (n 62) 3.

[71] Midgley, 'Manifesto' (n 68).

[72] Murdoch (n 48) 7-8 (emphasis added). This part of the book first appeared as an essay in 1964.

[73] *Ibid* 9.

[74] Susan Dodds, 'Choice and Control in Feminist Bioethics' in Catriona Mackenzie and Natalie Stoljar (eds), *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self* (Oxford University Press 2000) 213, 216.

chapter, want a better version of autonomy. She is clear that this view of persons 'is not, *and cannot by tinkering be made*, the philosophy we need'.[75] Her position is common to the whole group. Their objection is not to a particular abstract model of selves, but to the very assumption that philosophy's job is to reduce the messy moral complexity of human lives to abstract models. Opposition to that deeper assumption is present in Anscombe's provocative recommendation that modern philosophy should not even attempt ethics until it had paid more attention to the philosophy of psychology.[76] It is there, too, in Midgley: 'the academic's dream of pure, sanitized objectivity only leads us to conceal essential material'.[77] Most visibly, it is there in Murdoch's constant reminders that knowing the good in everyday life is difficult because 'the world is aimless, chancy, and huge, and we are blinded by self'.[78] Relational autonomy theorists attempt to modify the concepts that philosophers use, while leaving the relationship between philosophy and the world untouched. The golden age position is more similar to Harrington's recent argument against the very idea of 'law straightforwardly implementing the reasoned dictates of ethics'.[79] Harrington, however, turns to the disciplines of rhetoric, history, and economics. The golden age philosophers, instead, developed a different way of doing ethics.

The critical work in this thesis can be cast as an example of a single golden age problematic: if ethics consists only of a handful of highly abstracted concepts, then it will be 'too unrealistic to be applied to real life'.[80] Chapter two examines Christman's work precisely because it is more practically-orientated than other dual models, yet finds it cannot meaningfully guide practice. Similarly, the first part of this chapter surveys alternative models of autonomy and finds that they, too, show little prospect of guiding practice. In other words, the problem is not with Christman's implementation. It is with the idea that a concept of autonomy can determine when it is appropriate to act against a person's expressed wishes. Self-government is too marginal a part of human lives, and those lives are too intertwined in the world, for it to serve any such purpose.

[75] Murdoch (n 48) 46 (emphasis added).
[76] GEM Anscombe, 'Modern Moral Philosophy' (1958) 33 Philosophy 1, 12.
[77] Midgley (n 1) 160.
[78] Murdoch (n 48) 97.
[79] John Harrington, *Towards a Rhetoric of Medical Law* (Taylor and Francis 2017) 2.
[80] Midgley, 'Manifesto' (n 68).

Murdoch and her contemporaries would not allow a diagnosis like this to serve as an excuse to avoid doing ethics. After all, 'the search for unity is deeply natural' so 'moral philosophy ought to be defended'.[81] Instead, they take another approach to ethics, a ground-up and multifaceted one. Midgley summarises as follows:

> I and my friends did not try to claim credit for introducing any beautiful new simplicity. Far from that, we rather emphasized that these matters are really difficult and complicated – that we do indeed seriously need to think harder about them, so as to evolve concepts that will fill in the vast blank spaces that have been allowed to accumulate around the narrow ranges of our own experience. In fact, we all need to do some serious philosophizing here. And we ourselves have tried to suggest ways in which this could be done.[82]

There is a great deal here for anyone who takes the ethics of mental capacity law seriously. Indeed, if the critical work in this thesis demonstrates a particular example of the group's general diagnosis, then the positive work is an example of their recommended response. First, as stressed at the end of the last chapter, is the acknowledgment 'that these matters are really difficult and complicated'.[83] This point also guides the next two chapters. Chapter six shows that assessing a person's capacity entangles the assessor in a paradox of knowledge. It is possible to navigate this paradox, but doing so suggests moral duties to the person assessed that domestic law does not require an assessor to respect. Chapter seven, similarly, suggests the binary between autonomy and protection that often emerges in cases concerning the abuse of an adult with mental disabilities must be supplemented by a concept of humility. Doing this, however, only emphasises how difficult it can be to respond to these situations well.[84]

The other half of Midgley's equation, the need 'to evolve concepts that will fill in the vast blank spaces',[85] has also guided this thesis. Sometimes, attention to practice alone is enough to reveal conceptual complexity. The complex intertwining of capacity, voluntariness,

---

[81] Murdoch (n 48) 74.
[82] Midgley, Manifesto II (n 68).
[83] *Ibid.*
[84] This dynamic may be a general feature of a ground-up approach to morality. Attention to real moral difficulties requires the development of new concepts to navigate them; but the more moral concepts that guide a person's attention, the more real moral difficulties they will recognise.
[85] *Ibid.*

authenticity, intelligibility, and subjectivity discussed in chapter three emerged organically from the case reports. Similarly, the importance of false belief, discussed in chapter four, became obvious when capacity assessments were compared to dual models. Sometimes, however, attention to practice is not enough. Sometimes, concepts must be deliberately developed; and here, too, much can be learned from the golden age. One simple technique is to use examples. The wartime group did this so effectively that many of their examples have become standard. Anscombe's story of a man who is moving his hand up and down, which is pumping water, which is poisoning the people in a house, has become the classic example of the point that every intentional action occurs under multiple descriptions;[86] and Foot's example of a man on a runaway trolley who can only steer between two tracks, one of which will kill one person and one of which will kill five,[87] has become the basis of an entire field of psychological research.[88] There is, however, an irony to the success of these examples. They are arresting because they were formulated, in part, to break up the abstraction of philosophy, but their 'ritual' repetition has tended to make the examples themselves abstracted.[89] This thesis, as must be clear by this point, uses examples to interrogate philosophical ideas. It does so in two ways. Imaginary examples, for example my unreflective love for chips in chapter two or the difference between Ludwig and Winston in chapter four, show areas where a theory does not seem to fit every situation. They are essentially destructive. In contrast, real examples, primarily from reported cases, show what is done in practice, in order to gain a fuller picture of the sorts of circumstances that theory must be responsive to. For instance, chapter three interrogates the various ways that judges talk about the relationship between autonomy and capacity, and this points towards something more complicated than the gatekeeper account.

It is unlikely that examples alone can 'evolve concepts that will fill in the vast blank spaces',[90] but the wartime group also uses another technique. As Midgley says, 'when the attempt to look forward merely confuses us, we can often get a sudden light by looking

[86] GEM Anscombe, *Intention* (first published 1957, 2nd ed, Harvard University Press 2000) §23.
[87] Philippa Foot, 'The Problem of Abortion and the Doctrine of the Double Effect' (1967) 5 Oxford Review 1.
[88] Christopher W Bauman and others, 'Revisiting External Validity: Concerns about Trolley Problems and Other Sacrificial Dilemmas in Moral Psychology' (2014) 8 Social and Personality Psychology Compass 536 gives a critical review.
[89] Midgley, Manifesto (n 68); *ibid*.
[90] Midgley, Manifesto II (n 68).

back'[91] Often, they look back to classical Greece. Anscombe almost singlehandedly provoked the post-war revival of interest in Aristotle's ethics,[92] and Foot's *Natural Goodness* is a classic defence of this broad position.[93] Similarly, Murdoch is one of the few modern defenders of a Platonic conception of the Good (the capital is hers) as a real object: always beyond us, yet always drawing us towards it.[94] Their 'looking back', however, must be understand in the broader context of their 'multi-faceted programme'.[95] It is not an attempt to replace fashionable theories with older ones. They are suspicious of the very 'idea of a single fundamental explanation',[96] so they do not look to the past for one. For instance, the group has been influential in the recent renaissance of virtue terms, such as 'courage' or 'humility'; but their work is not 'virtue ethics' in the narrow sense of an attempt to reduce all ethics to virtue.[97] 'Looking back' simply acknowledges that people have always tried to understand an old Socratic problem, 'the way we ought to live',[98] and that there is little reason to think that only those close to us in time will have parts of the answer. In other words, if the group are traditionalists, then it is in Chesterton's sense of tradition as a 'democracy of the dead' that 'refuses to submit to the small and arrogant oligarchy of those who merely happen to be walking about'.[99] This thesis emulates the wartime group's work by drawing directly on historical works. Aristotle and Kierkegaard have already appeared

[91] Midgley (n 1) 183.
[92] Anscombe (n 76).
[93] Philippa Foot, *Natural Goodness* (Oxford University Press 2001).
[94] Murdoch (n 48) 89-97; Nora Hämäläinen, 'What is a Wittgensteinian Neo-Platonist? – Iris Murdoch, Metaphysics and Metaphor (2014) 43 Philosophical Papers 191.
[95] Wiseman (n 69).
[96] Midgley (n 1) 43. Murdoch insists that there is a unity to morality, but doubts that we can know it (n 48) 97. Similarly, Foot argues that happiness, good, and virtue are 'conceptually inseparable' and cannot be reduced to one another (n 93) 94. Anscombe is scathing about overly abstract ethics, observing that 'young men at Oxford …learned more definite, new moral theory' from the Nuremberg trials than from 'any teaching of moral philosophy': GEM Anscombe, 'Does Oxford Moral Philosophy Corrupt Youth?' (1957) 57 The Listener 266.
[97] Yang Xiao, 'Virtue Ethics as Political Philosophy: The Structure of Ethical Theory in Early Chinese Philosophy' in Lorraine Besser-Jones and Michael Slote (eds), *The Routledge Companion to Virtue Ethics* (Routledge 2015) 471 distinguishes between reductionist 'virtue ethics' and 'theories of virtue' that simply treat virtue as an important part of moral thought. Golden age philosophers have generally engaged in the latter.
[98] Plato, *Republic* in John M Cooper (ed), *Plato: Complete Works* (first published approximately 380BC, GMA Grube and CDC Reeve tr, Hackett 1997) 971, 352d.
[99] GK Chesterton, *Orthodoxy* in David Dooley (ed), *GK Chesterton: Collected Works Volume I* (first published 1908, Ignatius Press 1986) 209, 251. This attitude to time can now also be taken to geography. The last few decades have seen an explosion of detailed work in the English language on non-European philosophical traditions, from the Chinese to the Native American. The assumption that only European thinkers are relevant when addressing any aspect of the Socratic problem now looks as indefensible as the assumption that only recent thinkers are relevant: Bryan W Van Norden, *Taking Back Philosophy: A Multicultural Manifesto* (Columbia University Press 2017).

in passing; but the next two chapters draw on the past in a more systematic way. Chapter six uses Meno's paradox, which appears in one of Plato's dialogues, to interrogate the epistemology of mental capacity assessments; and chapter seven uses the concept of humility, largely in Aquinas's formulation, to fill some of the blank spaces that are left when the ethics of responding to the abuse of people with mental disabilities is reduced to a binary between autonomy and protection.

Before introducing these next two chapters, there is a final feature of the golden age group worth mentioning. Their writing is accessible. It is subtle without using jargon, and erudite without assuming that the reader is familiar with the literature. This is profoundly important to the ethics of mental capacity law, where philosophical, clinical, and legal jargons often collude to render the area obscure to outsiders. When that happens, the people being talked about, whose lives are affected by these disciplines, stand no chance at all of being included in the conversation. This thesis is not as well written as the wartime group's work. The author is not that skilful. Its core ideas, however, especially 'false belief' and 'humility', are not jargon. They are abstractions, but of the sort that people use in everyday life. They have 'a degree of precision', as Aristotle says, 'appropriate to the underlying material'.[100] If ethics is to connect to the world, then its style is as important as its content. The golden age of female philosophy provides an exemplar of both.

3.3 Two papers towards conceptual plurality

The next two chapters are published articles, appearing in Medical Law Review and the International Journal of Law and Psychiatry respectively. Each is written from the perspective developed to this point. They both assume the following: that actions against a person's expressed wishes must be justified in terms of the person's relationship to the world, not their relationship to themselves; that there is no existing 'model of autonomy' that can meaningfully settle when it is permissible to act against a person's expressed wishes; that questions about the permissibility of such actions are nevertheless unavoidable and important, so it is necessary to draw on other moral concepts; and finally, that these other concepts do exist, but each can only illuminate a small piece of the overall picture.

[100] Aristotle, *Nicomachean Ethics* (first published approximately 322BC, Christopher Rowe tr, Oxford University Press 2002) 1094b13.

Within this framework, each chapter addresses in more detail areas that have already been touched upon.

The next chapter, 'The Implications of Meno's Paradox for the Mental Capacity Act 2005', is, in some ways, an elaboration of one aspect of the 'necessity' step of the four-stage justification. In the last chapter, it was noted that necessity had a 'silent internal clause': an action against someone's expressed wishes is only necessary if they cannot be persuaded to change their false belief.[101] This leads directly to a version of a paradox that occurs in Plato's *Meno*.[102] Does the person believe that their belief is false? If they believe their first-order belief is false, then they do not believe it.[103] If they do not believe it is false, then they have no motivation to change it. The chapter adopts terminology from the *Meno's* secondary literature, so it casts this in terms of recognition. It notes that 'it is not enough that someone fail to recognise a need; they must also fail to recognise that they have failed to recognise the need'.[104] In the language of false belief, it is not enough that someone has a false belief about the world, they must also falsely believe that belief to be true. This may seem an abstract point, but concrete, under-examined, duties follow from it. First, it is necessary to take the person's own testimony as to whether or not they have capacity seriously; and, second, after acting against a person's wishes there is a duty to help them avoid false beliefs of this type in the future. In this respect, the chapter illustrates a point stressed at the end of the previous chapter. Basing the justification for actions against a person's expressed wishes on their relationship to the world, not their relationship to themselves, makes every single act of justification more difficult. This is a good thing.

Chapter seven, 'Humility when Responding to the Abuse of Adults with Mental Disabilities', is an example of approaching the law guided by the golden age of female philosophy. Situations of abuse are often reduced to a clash between imperatives to 'protect' and 'respect autonomy', but the chapter shows that this reduction eliminates from view the third party who faces the questions of whether and how to intervene. Humility, in contrast, makes the potential responder a direct object of evaluation. The chapter draws

[101] Chapter four, subsection 3.3.
[102] Plato, *Meno* in John M Cooper (ed), *Plato: Complete Works* (first published approximately 385BC, GMA Grube tr, Hackett 1997) 870, 80d80e.
[103] Ludwig Wittgenstein, 'Philosophy of Psychology – A Fragment' in PMS Hacker and Joachim Schlute (eds), *Philosophical Investigations* (first published 1953, GEM Anscombe and others tr, Wiley-Blackwell 2009) §91-92.
[104] Chapter six, section 2.

directly on Murdoch's work, but it also draws indirectly on the wartime group. Dusting off an old, slightly neglected, concept is one of their characteristic moves; and, more fundamentally, the chapter, like their work, assumes that that successfully navigating moral difficulties is not simply a matter of finding the right action, which can be determined in the abstract. It is also a matter of paying attention to the right things. In these cases, professionals must attend to their own limits. Finally, the chapter does not offer humility as a theoretically elegant solution to problems, just as one part of a larger picture. Chapter four called that larger picture 'truth' and noted that, for their wishes to be respected, the four-stage justification required people have some commitment to it. Chapter seven gives the other side of that coin. Humility requires that anyone relying on the four-stage justification to act against another person's expressed wishes also be committed to truth.

4. Conclusion

This chapter has two main sections. Section two argues that relational alternatives to value-neutral dual model autonomy, whether substantive dual models or pure social views, offer no prospect of distinguishing between when it is and is not permissible to act against a person's expressed wishes. In these circumstances, it seems likely that no model of autonomy can do this work. Section three goes further, and argues that an approach to ethics that reduces it to a handful of elegant concepts is of no use when confronting the situations faced by mental capacity law. Live alternatives, however, exist. The golden age of female philosophy developed several ways to move beyond ethical abstraction: using examples to draw attention to everyday situations; rescuing neglected concepts from the dusty attic of philosophical history; and writing in plain English. Finally, this chapter introduced the next two, which are published papers, and situated them within the thesis as a whole. They show how orientating justification to truth, not to autonomy, does not weaken respect for a person's expressed wishes. It strengthens it.

Chapter Six

The implications of Meno's paradox for the Mental Capacity Act 2005

Published paper: (2016) 24 Medical Law Review 279

1. Introduction

When, if ever, is it right to make a decision for another person that they object to? The question seems to be intractable. The Mental Capacity Act 2005 ('the MCA') provides one set of answers; but the UN Convention on the Rights of Persons with Disabilities ('the UNCRPD')[1] offers another, apparently incompatible, set.[2] The MCA says that such a decision can be made if it is in someone's best interests and they cannot understand, retain, or use and weigh the relevant information due to 'an impairment of, or a disturbance in the functioning of, the mind or brain'.[3] The UNCRPD, in contrast, has been interpreted to prohibit all 'substituted decision-making': defined as any process in which a third party is appointed against someone's will to make decisions for them; and the decision is made in the person's best interests, rather than according to the best interpretation of their will and preferences.[4] Compelling points have been made against both of these positions, and each has been given increasingly subtle interpretations to accommodate these objections.[5] The

* This work was first presented at 'The Mental Capacity Act 2005: Ten Years On' Conference at Liverpool on 9-10 September 2015, which was supported by the Wellcome Trust (Grant No 107461/Z/15/Z).

[1] Convention on the Rights of Persons with Disabilities (adopted 13 December 2006, entered into force 3 May 2008) 2515 UNTS 3.

[2] Peter Bartlett, 'The United Nations Convention on the Rights of Persons with Disabilities and Mental Health Law' (2012) 75 Modern Law Review 752.

[3] Mental Capacity Act 2005 s1-4. An inability to communicate, under s3(1)(d), is not relevant to a decision that someone objects to.

[4] UN Committee on the Rights of Persons with Disabilities, 'General Comment No 1' (2014) UN Doc CRPD/C/GC/1 para 27.

[5] Elionóir Flynn and Anna Arstein-Kerslake, 'The Support Model of Legal Capacity: Fact, Fiction, or Fantasy?' (2014) 32 Berkeley Journal of International Law 124; Wayne Martin and others, 'Achieving CRPD Compliance' (Essex Autonomy Project 2014) 10-13; John Dawson, 'A Realistic Approach to Assessing Mental Health Laws' Compliance with the UNCRPD' (2015) 40 International Journal of Law and Psychiatry 70; Lucy Series, 'Relationships, Autonomy and Legal Capacity: Mental Capacity and Support Paradigms' (2015) 40 International Journal of Law and Psychiatry 80; Piers Gooding, 'Navigating the 'Flashing Amber Lights' of the Right to Legal Capacity in the UNCRPD: Responding to Major Concerns' (2015) 15 Human Rights Law Review 45.

resulting dialogue, however, does little to dispel the impression that the core question is intractable; for, on this, neither side has appreciably changed its answer.

This chapter does not solve the question of whether it is better to make decisions for others in their bests interests or according to the best interpretation of their will and preferences. Instead, it draws attention to a paradox that can occur in supportive relationships. From this paradox, the chapter shows that best interests decision-making against a person's will, *if* it is justified in a liberal society, must be premised on the decided-for person being unable to recognise that they have failed to recognise a need. From this premise, two duties can be derived; and this allows any particular system that uses a best interests standard, in this case the MCA, to be evaluated against standards that it can be assumed to be implicitly committed to. This is not the same task as assessing the Act against an external normative framework, such as the one that the UNCRPD presents. Nevertheless, it is striking that if these duties were taken seriously, then it would narrow, without entirely closing, the controversial gap between the two systems.

The next section develops an account of the paradox. Section 3 derives the first duty, the duty of identification, and assesses the MCA against it. Section 4 then evaluates the extent to which legal requirements to provide advocacy might require compliance with this duty. Finally, Section 5 derives the second duty, the duty to support, and assesses the Act against this. For each duty, it is found that the Mental Capacity Act allows compliance, but stops some way short of requiring it. In the conclusion, these limited criticisms of the Act are reintegrated into the context of the wider debate.

2. Meno's paradox in supportive relationships

Meno's paradox is presented by Plato in the dialogue of the same name. It is stated in two ways, first by Meno and then by Socrates.[6] Socrates' statement of the problem is slightly clearer. He says:

[6] Plato, *Meno* in John M Cooper (ed), *Plato: Complete Works* (first published approximately 385BC, GMA Grube tr, Hackett 1997) 870, 80d – 80e.

> [A] man cannot search either for what he knows or what he does not know… He cannot search for what he knows – since he knows it, there is no need to search – nor for what he does not know, for he does not know what to look for.[7]

A common response to this paradox is to deny that knowledge is as binary as it suggests. So, for instance, Fine interprets Plato's Socrates as proposing a distinction between 'knowledge' and 'true beliefs'.[8] On this interpretation, we may believe things to be true that are actually true without knowing why they are true.[9] These true beliefs then provide us with a starting point from which we can develop a deeper understanding;[10] and this understanding may, in time, develop to the point where we can give a coherent account of why our beliefs are true. It is this ability to explain 'why' that distinguishes knowledge from belief.[11] There is, however, a form of the paradox that this distinction does not resolve.

Meno originally asks Socrates 'how will you know it is the thing you didn't know?'[12] This suggests a particularly difficult problem, the problem of recognition.[13] If knowledge requires being able to give an account of the reasons why a belief is true, then those 'reasons' will inevitably refer to other beliefs that, to count as knowledge, must in turn rely on even deeper reasons. There is no obvious end to this process, and so no certain way of distinguishing between what you know and what you falsely believe that you know.[14] There is always the possibility that something we take to be knowledge is actually premised on a deep unrecognised error. Sometimes, this is directly relevant to supportive relationships.

A paradox of recognition can occur in supportive relationships. This can be shown by an example. Imagine a man called 'John', who has a large (over 9cm) abdominal aortic aneurysm that is growing rapidly. John wants to live. The diagnosis of such an aneurysm is reliable, they are very likely to rupture, and rupture carries an extremely high risk of

---

[7] *ibid* 80e.
[8] Gail Fine, 'Inquiry in the Meno' in Richard Kraut (ed), *The Cambridge Companion to Plato* (Cambridge University Press 1992) 200, 206.
[9] Perhaps because we simply accept what an authority, such as a teacher, says. Alasdair MacIntyre, *Three Rival Versions of Moral Enquiry* (Duckworth 1990) 63.
[10] Scott distinguishes between knowing 'parrot fashion' and 'synoptically'. Dominic Scott, *Plato's Meno* (Cambridge University Press 2006) 102-103, 183-184.
[11] Plato (n 6) 98.
[12] *Ibid* 80d.
[13] Nicholas P White, *Plato on Knowledge and Reality* (Hackett 1976) 42-47.
[14] Scott (n 10) 83-84.

death.[15] Assuming, for now, that there are no other complicating factors, it is reasonable to say that John objectively needs a particular form of support: surgery. Further imagine, however, that John does not recognise his need for this support. He is asymptomatic, and simply disbelieves the diagnosis. In these circumstances, those supporting him might want to help him to recognise that, if he wishes to live, he needs surgery. This is when the paradox appears. If John accepts this help – a second-order support to recognise his need for first-order support – then that may give us a reason to suspect that he is not as certain as he seemed about not needing surgery. To accept the second-order support shows some recognition of the fact that he may have the first-order need, surgery. This is the paradox; and, more importantly, it also works in the opposite direction. If John truly does not recognise his need for surgery, then he will not recognise that he has a reason to accept help to recognise that need. If a supporter says, 'we want to help you to see that you need surgery', then he is likely to reply 'but I don't need surgery'. Unless he is a particularly patient man, this might be followed by a comment about how nobody is listening to him.

To this use of the paradox it might be objected that John could recognise his need for surgery yet refuse it. This raises two possibilities. An example of the first would be if John wished to live, and understood that he needed surgery to do so, but nevertheless decided that a religious prohibition made it impossible for him to accept the intervention. In this case, however, there is good reason to doubt whether John truly needs the surgery, no matter what his medical needs might be.[16] In contrast, an example of the second possibility would be if John wanted to live, and accepted that he needed surgery to live, but nevertheless refused the intervention without being able to give any reason why. In this case, however, what has occurred is not recognition. The concern in these situations is with a practical question, not with someone's ability to assent to theoretical propositions; so 'recognition' here necessarily includes the ability to make practical use of what has been recognised. This roughly accords with the natural use of the language. When, in this example, John agrees he needs surgery to live he has clearly assented to a proposition. After he refuses the surgery for no discernible reason, however, it would be unusual to nevertheless say that he 'recognises' his need for surgery. In conclusion, in the first example

[15] FL Moll and others, 'Management of Abdominal Aortic Aneurysms: Clinical Practice Guidelines of the European Society for Vascular Surgery' (2011) 41 European Journal of Vascular and Endovascular Surgery S1.
[16] This point is dealt with in detail in the next section, but for now it is enough to observe that someone's objective needs cannot be identified without taking into account their own understanding of their life.

John recognises an apparent need but denies that in the circumstances it is a need for him; and in the second what appears to be recognition is better characterised as mere assent. In neither case has John recognised a need for him and yet refused surgery.

Examining this paradox can help to reveal one condition necessary to justify systems, such as the MCA, that allow decisions to be made in someone's best interests against their will. For such decisions to be justified, it is not enough that someone fail to recognise a need; they must also fail to recognise that they have failed to recognise the need. If a person can recognise their own first-order needs, then they are capable of recognising their own best interests; so, in a liberal society, there is no obvious justification for giving someone else the power to make such decisions for them. Beyond this, however, if someone does not recognise a first-order need but knows that they do not do so, then there is still no clear justification for giving someone else the ability to make decisions for them *against their will*. If, for example, John says 'I must be missing something, you decide for me', then the decision is not against his will. If, in contrast, he says 'I must be missing something, please explain it to me again', then there is no clear justification for ignoring his request for support and simply making the decision. When someone recognises their second-order need for help to identify their first-order needs, then less intrusive measures than a decision against their will are necessarily available. In other words, if it is assumed that self-determination is generally good for people, and that any legal suspension of it should be for a reason, then the mere fact that people sometimes have needs that they cannot recognise is not enough to justify best interests decisions against someone's will. If such decisions are justified, it must be because sometimes people fail to make the second-order recognition that they cannot recognise a first-order need. That is to say, systems that allow best interests decisions against a person's will are only justified if people are sometimes caught in the paradox.

The argument in the previous paragraph does not demonstrate that best interests decisions against a person's will are ultimately justified. It may be that even if someone is caught in the paradox, then making a decision against their will is still unjustified for other reasons. For example, it might always cause significantly more harm to the person than respecting their preferences would. For this reason, reflecting on the paradox cannot settle the clash between the MCA and the UNCRPD discussed in the introduction. Nevertheless, if a second-order failure of recognition is necessary for a best interests decision against

someone's will to be justified in a liberal society, then there are implications. In particular, the power to make such decisions will plausibly carry with it duties.

3. The duty of identification

The previous section suggests that if best interests decisions made against someone's will are to be justified in a liberal society, then it is not enough that the person does not recognise a first-order need; they must also fail to recognise that they do not recognise a first-order need. From this general form of the argument, a specific form can be derived: in any particular case, a best interests decision against a person's will is only justified if they both fail to recognise a first-order need and fail to recognise that they are doing so. If it is assumed that decisions should only be made when it is justified to do so, then from this specific form of the argument a duty can be derived: the duty to identify whether someone is making such a second-order failure of recognition before deciding against their will.

There is, however, a problem; for this duty can become ensnared in the paradox that it is derived from. If the duty is fulfilled, a decision is only made when the person cannot be helped to make the second-order recognition that they have failed to recognise a first-order need. If, though, the person, like John in the original example, is in this position, then they will deny that this is an accurate description of their situation. In other words, the same conditions that are necessary to justify the power to make a decision also make the use of that power unlikely to be accepted. This has profound consequences. So far, this chapter has largely taken a need as a simple objective fact; but, of course, no supporter will ever have unmediated access to the objective facts. They will have, at best, a tangle of evidence from various sources about what someone's needs are. Furthermore, even a person's objective needs will be influenced by subjective factors. Even in the example, it was necessary to say that John wanted to live. A person's experiences and values partially shape their needs. This creates an evidential problem, for each of us has an access to these subjective elements that an assessor cannot equal. Atkins puts the point well: 'appreciation of the subjective character of experience brings with it the necessity for an epistemological humility' on the part of others.[17] If the supported person has better access to some of the evidence, then there is always the possibility that what appears to be a case of them failing

[17] Kim Atkins, 'Autonomy and the Subjective Character of Experience' (2000) 17 Journal of Applied Philosophy 71.

to recognise a need is actually a case of third parties attributing needs to them that they do not have.[18]

Subjective elements shape our needs, so it will often be difficult to show that someone has needs that they are failing to recognise. This does not, however, mean that providing sufficient evidence is always impossible. Unfortunately, the UN Committee on the Rights of Persons with Disabilities appears to have concluded otherwise. It says that it is a flaw to presume to be able to 'accurately assess the inner-workings of the human mind'.[19] If this is a flaw it is, as Dawson points out, one that can be found throughout the legal system;[20] but, beyond that, the Committee fails to apply its own position consistently. Some evidence is only accessible to the supported person; but this is seldom the only relevant evidence that is only accessible to one person. Imagine John says that he doesn't need surgery, and supports his claim by saying that the doctors have kidnapped him in order to harvest his organs. In this case the staff, by virtue of *their* subjective experience, will have access to evidence that is not available to John. Some of this evidence, for instance the doctor's intention to save John's life or her memories of learning how to interpret the relevant scans, will bear on the question of what John's needs actually are. The existence of subjective perspectives does not, by itself, call into question the existence of objective needs. It just means that these situations are evidentially complex.

Evidentially complex situations are not unusual. As Wimsatt points out, 'for the complex systems encountered in evolutionary biology and the social sciences, it is often unclear what is fundamental or trustworthy'.[21] A common response is to seek robust results: if different types of measurement all return the same result, then that result can be treated as more trustworthy.[22] This idea of seeking robust results is relevant here; for if the person concerned has access to evidence that cannot be gained in any other way, then this implies that their own assessment of the situation should never simply be ignored. This in turn suggests that the duty of identification is only fulfilled if the person's own perspective on

[18] Judges in the Court of Protection are aware of this danger. See, for example, *V v Associated Newspapers* [2016] EWCOP 21, [2016] COPLR 236 [67]-[68] (Charles J).
[19] General Comment (n 4) para 15.
[20] Dawson (n 5) 74.
[21] William C Wimsatt, *Re-engineering Philosophy for Limited Beings* (Harvard University Press 2007) 56.
[22] *Ibid* ch 4. This process is not infallible. 'Robust detection' will break down if apparently independent measurements are not truly independent. Brett Calcott, 'Wimsatt and the Robustness Family' (2011) 26 Biology and Philosophy 284.

their needs is sought out and treated as direct evidence of whatever it is that they assert. Only if the countervailing evidence is overwhelming should the person's view be treated as evidence that they cannot recognise an objectively existing need. In practice, this may equate to Banner and Szmukler's application of Davidson's Principle of Charity: 'we must assume that a speaker is by and large consistent and correct in his beliefs'.[23] Indeed, if a supporter takes the opposite approach, and *begins* by taking a person's disagreement as evidence that they cannot identify their needs, then it is not the person who does not recognise that they have failed to recognise something. It is the supporter.[24] The approach suggested here will not, of course, resolve every possible evidential problem. It cannot address the question of what should first trigger the suspicion that someone does not recognise their own needs. It can only suggest what should be done when that suspicion is already present. Similarly, it cannot help if it is suspected that someone cannot recognise a need; but, despite every available support, they cannot or will not share their own perspective. This last point can, however, be accommodated by acknowledging that the duty of identification requires starting with the person's own perspective only when it is possible to do so.

If systems which allow best interests decisions against a person's will should respect the duty of identification, and the duty of identification requires someone's testimony to be treated as evidence of what they say before it is treated as evidence that they do not understand, then this offers two criteria by which the Mental Capacity Act can be assessed. The first criterion is whether it recognises the duty of identification at all, and the second is whether it recognises that the duty entails this careful treatment of the person's own testimony. It meets the first criterion. Capacity assessments under the MCA are of someone's ability to make a particular decision;[25] and an inability to make a decision is specified as an inability to understand, retain, use, or weigh relevant information; or an inability to communicate the decision.[26] The statutory language is wider than that derived from the paradox; but 'relevant information' includes the consequences of the decision,[27]

[23] Donald Davidson, 'Psychology as Philosophy', in *Essays on Actions and Events* (Oxford University Press 2001) 229, 238; cited in Natalie F Banner and George Szmukler, '"Radical Interpretation" and the Assessment of Decision-Making Capacity (2013) 30 Journal of Applied Philosophy 379.
[24] Thanks to David Gibson for this point.
[25] MCA s2(1); *PC* v *City of York Council* [2013] EWCA Civ 478, [2014] 2 WLR 1 [35] (McFarlane LJ).
[26] MCA s3(1).
[27] MCA s3(4).

and this seems close to being unable to identify your needs in that particular situation. This impression is reinforced when the Act's principles are examined. Section 1(2) requires that someone is assumed to have capacity unless the contrary is established. If being unable to understand or use information is analogous with the content of the duty of identification, then it is this section of the Act that makes it a duty to undertake a process of identification in every case. Furthermore, this principle is reinforced elsewhere in the Act. Section 5(1) requires that, for acts 'in connection with' care or treatment, 'reasonable steps' are taken to establish whether the person has capacity; and section 1(3) requires 'all practicable steps' be taken to help someone to make a decision before they are treated as unable to do so. This last section should distinguish between those who can be helped to recognise their needs and those who cannot, and only allow decisions in the latter case. The Act, especially this provision, has 'not been widely implemented';[28] but it certainly recognises the existence of something broadly similar to the duty of identification. Unfortunately, the same cannot be said of the second criterion, the careful treatment of the person's own testimony.

The Act's near silence on the question of evidence prevents it from requiring that the person's own testimony be treated with the care that the duty of identification demands. It does contain a few evidential prohibitions: incapacity cannot be established merely because of an unwise decision; or because of a person's condition, behaviour, age, or appearance.[29] It does not, however, indicate how the person's own perspective should be treated. The Code of Practice examines capacity assessments in more detail than the Act;[30] but it does not mention the person's perspective in this context. Although it discusses the possibility of challenging a capacity assessment,[31] the standard against which the assessment will be measured is that of the Act and Code;[32] so this takes things no further. Worse yet, an ambiguity in the language of the Code may mislead. It says that 'nobody can be forced to undergo an assessment of capacity'.[33] At face value, this would appear to mean that people

[28] Select Committee on the Mental Capacity Act 2005, *Mental Capacity Act 2005: Post-Legislative Scrutiny* (2013-14, HL Paper 139) paras 79-83, 103-110.
[29] MCA s1(4), 2(3). The word 'merely' is significant – these factors can have evidential force if they do not individually determine the question – *D* v *R (the Deputy of S)* [2010] EWCOP 2405, [2011] WTLR 449 [40] (Henderson J).
[30] Department for Constitutional Affairs, *The Mental Capacity Act 2005 Code of Practice* (The Stationery Office 2007) ch 4.
[31] *Ibid* paras 4.63–4.65.
[32] *Ibid* para 4.64.
[33] *Ibid* para 4.59.

have the ability to veto their own capacity assessment; and, although framed in an unhelpfully confrontational way, this would mean that their perspective would have to be taken into account. The Code has not been read in this way, as the 'DD' series of cases illustrate. In these cases, there was consensus that 'threats or attempts to force DD to agree to an assessment are not acceptable',[34] but it was nevertheless declared lawful to remove her from her home 'by force' to a place where assessment could take place.[35] Her lack of cooperation with the subsequent assessment was respected;[36] but, importantly, this did not prevent her from being found to lack capacity to make decisions about litigation, contraception, and sterilisation.[37] In other words, DD could decline to take part in her capacity assessment; but that assessment took place without her anyway. There is no right to veto your own capacity assessment, and the Act contains no less extreme way to ensure that the person's own testimony is given a fair hearing, so the MCA fails to require that the second criterion of the duty of identification is fulfilled.

It seems likely that the Act's failure to ensure that someone's own testimony is treated as evidence contributes to such testimony being discounted in practice. Empirical studies show that a disagreement between an assessor and the person assessed can sometimes be taken as evidence that the person 'lacks insight', without first taking the disagreement as a reason to question the validity of the assessment.[38] A lack of insight appears to be strongly correlated with a finding of incapacity,[39] so the uncritical use of this concept is likely to have consequences. Professionals using 'insight' uncritically have been corrected by the Court of Protection,[40] but judges, too, sometimes seem to use the concept in an indiscriminate way.[41] When this happens, a person's disagreement can be discounted if they lack insight, and

[34] *The Mental Health Trust* v *DD (no 2)* [2014] EWCOP 13, 142 BMLR 156 [36] (Cobb J).
[35] *Ibid* [37]-[41].
[36] *The Mental Health Trust* v *DD (no 3)* [2014] EWCOP 44 [5], [8]; *The Mental Health Trust* v *DD (no 4)* [2015] EWCOP 4 [135] (Cobb J).
[37] *DD (n 3)* (n 36) [13]; *DD (n 4)* (n 36) [63], [80].
[38] Val Williams and others, *Making Best Interests Decisions: People and Processes* (Mental Health Foundation 2012), 55-58; Charlotte Emmett, 'Homeward Bound or Bound for a Home? Assessing the Capacity of Dementia Patients to Make Decisions about Hospital Discharge: Comparing Practice with Legal Standards' (2013) 36 International Journal of Law and Psychiatry 73, 77.
[39] GS Owen and others, 'Mental Capacity, Diagnosis and Insight in Psychiatric In-Patients: A Cross-Sectional Study' (2008) 39 Psychological Medicine 1389.
[40] *CC v KK* [2012] EWCOP 2136 [36],[64](Baker J).
[41] *London Borough of Islington* v *QR* [2014] EWCOP 26, (2014) 17 CCL Rep 344 [96] (Batten DJ); *An English Local Authority* v *SW* [2014] EWCOP 43, [2015] COPLR 29 [20] (Moylan J).

their disagreement is evidence that they lack insight;[42] so it becomes 'impossible to disbelieve a doctor and retain capacity'.[43] Here, the duty of identification has become entangled in the paradox that it was derived from; and it is the assessor, not the person assessed, who has failed to recognise what it is that they do not recognise. The Act allows this situation to be avoided. It does not, however, require that it be avoided.

4. Advocacy and the duty of identification

Since the MCA was passed, the ratification of the UNCRPD has made the perspective of the decided-for person a more prominent issue; and the effects of this new prominence can already be seen in domestic law, particularly in changes to the legal framework governing the provision of advocacy. These developments bring the law closer to requiring compliance with the second criterion of the duty of identification; but there are, however, limits to what advocacy can achieve in this respect. This section discusses the recent legislation, then its limits. First, however, it examines the advocacy provisions of the MCA.

The MCA requires an Independent Mental Capacity Advocate ('IMCA') to be appointed when someone is thought to lack capacity, and has no-one else to speak on their behalf; and faces serious medical treatment, a long term change of accommodation, or a deprivation of liberty.[44] IMCAs can challenge capacity decisions,[45] so they offer an avenue for challenging assessments which ignore the person's own testimony. Nevertheless, unless appointed under section 39D of the Act when someone is deprived of their liberty,[46] they are consulted to help determine what would be in the person's best interests.[47] This leads to an ambiguity. If the IMCA believes that the person lacks capacity, then they might not challenge an assessment that the person disagrees with; for they might not believe that it is in the person's best interests to do so.[48] In other words, IMCAs are merely allowed, not required, to argue for the person's own perspective.

[42] Kate Diesfeld and Stefan Sjöström, 'Interpretive Flexibility: Why Doesn't Insight Incite Controversy in Mental Health Law?' (2007) 25 Behavioural Sciences and the Law 85.
[43] Neil Allen, 'Is Capacity "In Sight"?'(2009) 19 Journal of Mental Health Law 165, 167.
[44] MCA s37-39E.
[45] The Mental Capacity Act 2005 (Independent Mental Capacity Advocates) (General) Regulations 2006, SI 2006/1832, reg 7(1)(b).
[46] *AJ v A Local Authority* [2015] EWCOP 5, (2015) 18 CCLR 158 [107-11].
[47] MCA s37(1)(b), 38(1), 39(1)(b), 39A(1)(b), 39C(1)(c).
[48] Select Committee (n 28) para 167.

The advocacy provisions in the Care Act 2014 are more robust than those in the MCA, and provide a useful contrast, as the trigger for advocacy is still framed in terms of ability to make a decision. The newer Act requires an advocate to be appointed during needs assessments, care planning, or safeguarding processes;[49] if 'the individual would experience substantial difficulty' making a decision, and no-one else can represent and support them.[50] This is a lower bar than in the MCA, for 'substantial difficulty' does not require the person to be thought to lack capacity;[51] but a more relevant difference between the two Acts is in what, exactly, the advocate is required to do once appointed. Under the Care Act, an advocate is *required* to communicate the 'views, wishes, or feelings' of a person thought to lack capacity;[52] and the statutory guidance reinforces this as a duty to put forward the person's own case.[53] This goes further than the MCA towards robustly fulfilling the second criterion of the duty of identification, and it may also influence the older Act. The Law Commission has suggested that IMCAs 'be replaced by a system of Care Act advocacy'; this would incorporate the more robust standard into the MCA itself.[54]

Strengthening the law governing the provision of advocacy is a step towards requiring compliance with the duty of identification, but it has limits. These limits are of three types: limits of scope, limits in implementation, and cultural limits. The legislation itself has limited scope. Only certain decisions trigger the duty to appoint an advocate. Between the two Acts, many major decisions will be covered, but it is not clear that all will be; and, as has been acknowledged in Parliament, 'minor' decisions can be as important to the person concerned as 'major' ones.[55] In practice, however, limitations in implementation are likely to be more severe than those of scope. Failures of implementation were the most consistent theme of the recent House of Lords Select Committee Report into the MCA,[56] and IMCAs are not always appointed even when it is legally required.[57] Legislation requiring advocacy

[49] Care Act 2014 s67(1), (3), 68(1).
[50] *Ibid* s67(4), 68(3).
[51] Department of Health, *Care and Support Statutory Guidance* (Department of Health 2014) paras 7.55, 7.58.
[52] The Care and Support (Independent Advocacy Support) Regulations 2014, SI 2014/2824, reg 5(7).
[53] Statutory Guidance (n 51) para 7.43.
[54] Law Commission, *Mental Capacity and Deprivation of Liberty: A Consultation Paper* (Consultation Paper No 222, 2015) para 9.41.
[55] Joint Committee on the Draft Mental Incapacity Bill, *First Report* (2002-03, HL Paper 189-I, HC 1083-I) para 27.
[56] Select Committee (n 28).
[57] Department of Health, *The Seventh Year of the Independent Mental Capacity Advocacy (IMCA) Service* (Department of Health 2015) 34.

of a certain level is not, by itself, sufficient to ensure that practice will consistently reach that level; and the current political and economic climate is relevant here. Local authority budgets have fallen by almost forty per cent in real terms over the last five years, and authorities have lost their ability to insulate social care from the effects of this drop in funding; yet the overall need for care is increasing dramatically.[58] In these circumstances, advocacy may be neglected; and the provision of more 'basic' physical care prioritised.

Scope and implementation offer relatively contingent limits to advocacy. Changes to the legislation, or to practice and funding, could overcome them. There are, however, also cultural limits to the ability of an advocate to fulfil the duty of identification, and these will be harder to overcome. These cultural limits gain their force from the paradox; but to show this requires careful disambiguation. The example of John can help. If an advocate wants to represent John's own perspective as faithfully as possible, then she faces a choice. John both wishes to live and wishes to refuse life-saving surgery. If she presents either of these wishes alone, then she will not fully represent John's relevant views; but if she presents both, then she will undermine John's case by emphasising its contradictions. In these cases, then, presenting what someone wants without explaining why they want it is insufficient. Instead, an advocate should present the person's overall view, for example John's view that refusing surgery will not risk his life, and the *reasons why* that view should prevail.[59]

If an advocate attempts to present the reasons why someone's view should prevail, then another choice reveals itself. This is the choice between only presenting John's reasons for his view and also presenting other reasons for reaching the same view. With regards to the duty of identification, these are different things. If the advocate presents John's reasons for his view, then she is helping to ensure that his perspective is treated as evidence, as the duty requires. If, however, she offers reasons which are not John's for why John's wishes should prevail, then a shift in argumentation has occurred. Proposing reasons that are not John's does not help to show that John understands the relevant information and hence has capacity. Instead, the advocate is effectively arguing that regardless of whether John has capacity or not, his wishes should prevail. This is no longer about the duty of identification;

[58] National Audit Office, *Care Act First-Phase Reforms* (2015-16, HC 82) paras 1.6 - 1.9.

[59] The same point is true of 'wishes and feelings' in best interests decisions: Laura Pritchard-Jones, '"This Man with Dementia"—'Othering' the Person with Dementia in the Court of Protection' (2016) 24 Medical Law Review 518.

it does not bear on the question of whether John recognises his needs. There are good reasons for keeping these two lines of argumentation distinct. Being found to have capacity is not the same as being found to lack capacity but nevertheless having your wishes followed. A finding of incapacity can carry costs to the person beyond simply not getting what they want: for instance, self-stigmatisation and loss of confidence.[60] Beyond this, it is only by disambiguating between the two that the cultural limits on advocacy can be made clear.

The distinction between advancing someone's own reasons for their view and presenting other reasons for why that view should prevail makes it easier to see the cultural limits on the ability of an advocate to fulfil the duty of identification. An advocate's position will often, regrettably, give her more credibility than the person themselves; but this credibility has limits, and some of these limits are cultural. If, again, John is paranoid and believes that he has been kidnapped so that his organs can be harvested, but is actually ill in hospital, then an advocate arguing that his reasons for refusing treatment are literally true will not help him. She will only harm her own credibility. For a reason to be convincing, it must have some coherence with the account of the world that listeners assume to be true. This is not, by itself, a problematic limit: in John's case, even a charitable treatment of his testimony should almost certainly conclude that he is failing to recognise a need. It does, however, indicate where the problematic cultural limits to advocacy are. Experimental psychology has shown the existence of implicit social biases. These biases are entirely distinct to someone's explicit attitudes, they are robust, and they predict behaviour.[61] In other words, people consistently act on biases that they deny, and may not even perceive. In particular, strong implicit biases against those with disabilities have been shown to be widespread, even among professionals who work closely with them.[62] These biases, combined with the need for a person's reasons to have some minimal coherence with commonly held assumptions to be credible, are sufficient to impose limits on advocacy's power to fulfil the duty of identification. Recognising the existence of widespread unacknowledged bias

[60] Bruce J Winick, 'The Side Effects of Incompetency Labelling and the Implications for Mental Health Law' (1995) 1 Psychology, Public Policy, and Law 6.

[61] Kristin A Lane and others, 'Implicit Social Cognition and Law' (2007) 3 Annual Review of Law and Social Science 427; Brian A Nosek and others, 'Implicit Social Cognition: From Measures to Mechanisms' (2011) 15 Trends in Cognitive Sciences 152.

[62] Michelle Clare Wilson and Katrina Scior, 'Attitudes towards Individuals with Disabilities as Measured by the Implicit Association Test: A Literature Review' (2014) 35 Research in Developmental Disabilities 294.

entails recognising the risk that sometimes when a person's reasons seem unconvincing, because they do not cohere with the commonly accepted view of the world, it is society, not the person, which has not recognised its failure to recognise significant facts. Furthermore, when the paradox reappears in this societal form, it limits advocates in two ways. First, advocates may be implicitly biased; for explicit opposition to discrimination does not guarantee that someone is not implicitly biased.[63] Second, if the advocate is not biased, and others are, then arguing for the person's perspective will risk the advocate's credibility in the same way, if seldom to the same extent, that arguing John really was going to have his organs harvested did. The advocate will literally be arguing against what 'everyone knows'. These cultural limits will be difficult to address: implicit biases may be difficult to change,[64] and such biases against those with disabilities do not seem to have fallen in the last decade.[65] They are, however, only limits. An area of freedom should remain in which an advocate can challenge biases without destroying their own credibility; and if that balance is difficult to find, then this is an argument for strong advocacy, not against it. Nevertheless, the cultural limits on advocacy, in combination with the limits of scope and in implementation, do mean that while stronger advocacy legislation is a step towards requiring that a person's testimony is treated with due care, it is only one step. Others will almost certainly be needed.

5. The duty to support

In addition to the duty of identification, a duty to provide support can be derived from the paradox; and, as with the first duty, this offers a criterion by which legal systems that permit best interests decisions against a person's will can be evaluated. At the outset, however, it should be stressed that 'support' in this context is limited to support of one particular, relatively narrow, type. It is support to someone who has had a decision made for them because of a second-order failure of recognition, which is directed at helping them to avoid such failures in the future. There are further limits on this support – in particular, it cannot be against the person's will – that are discussed below. First, however, this section shows how the duty follows from the paradox.

[63] Lane (n 61) 431.
[64] Jennifer A Joy-Gaba and Brian A Nosek, 'The Surprisingly Limited Malleability of Implicit Racial Evaluations' (2010) 41 Social Psychology 137.
[65] Wilson (n 62) 319.

As argued in sections 2 and 3, deciding against a person's will can only be justified if someone cannot recognise that they do not recognise a need. At the same time, however, interfering with someone's self-determination is understood in liberal societies to be wrongful. Indeed, this understanding motivates the general principles of both the MCA[66] and the UNCRPD.[67] On a simplistic reading, this general prohibition on interfering with someone's self-determination is entirely displaced when someone's second-order failure of recognition justifies making a decision against their will. Such a reading is, however, *too* simplistic; for it collapses the difference between doing an unqualifiedly right act and doing a justified wrong.[68] Making a decision against someone's will because of their second-order failure of recognition is not an unqualifiedly right act. The very structure of this description denies it. It presents an apparent wrong, making the decision against their will; and then presents a reason, the failure of recognition, for nevertheless doing it. It is, at best, a justified wrong; and this is important. Justified wrongs are necessarily cases in which moral conflicts have occurred and been resolved; but, as Williams observes, such conflicts are not 'all soluble without remainder'.[69] One 'remainder' he examines is the regret that an 'admirable moral agent' might feel, even when they are acting as well as circumstances seem to allow.[70] This is a good starting place; for it seems right to feel some regret if we have made a decision against someone's will, even if we are sure that doing so is justified in the particular case. Furthermore, it seems right that, as Williams says,[71] someone experiencing such regret might conclude that they ought to avoid such situations arising in the future. In the particular case, the moral conflict has arisen because someone has not recognised that they do not recognise a need; so a regretful decision-maker might conclude that, where possible, they should help the decided-for person to avoid such second-order failures of recognition in the future. This, then, is the derivation of the duty to support. If it is accepted that a decision against a person's will is not an unqualifiedly good act but a justified wrong, and that the appropriate response to inflicting this justified wrong is regret, and that the appropriate response to such regret is to attempt to avoid such situations occurring in the

[66] MCA s1.
[67] UNCRPD art 3(a).
[68] Rosalind Hursthouse, *On Virtue Ethics* (Oxford University Press 2001) 44-49 discusses this distinction.
[69] Bernard Williams, 'Ethical Consistency' in Geoffrey Sayre-McCord (ed), *Essays on Moral Realism* (Cornell University Press 1988) 41, 52.
[70] *Ibid* 49.
[71] *Ibid* 50.

future, then there is an apparent duty once a decision has been made against someone's will to attempt to help them to avoid second-order failures of recognition in the future. This is an endoxic derivation, it appeals to widely held evaluative beliefs;[72] so it can be rejected by someone who denies regret is an appropriate response to inflicting a justified wrong, or denies that we should act on such regret. Nevertheless, neither of those positions seem particularly appealing, so it does have some force. The duty to support, however, also faces an internal challenge.

Just like the duty of identification, the duty to support can become ensnared by the paradox that it is derived from. If the justification for making a decision against someone's will is that the person could not be helped to recognise their failure to recognise a need; then to claim that there is, nevertheless, a duty to provide this same help seems absurd, doomed to failure from the outset. This objection may hold true in some cases; and in those cases, if they can be reliably identified, there is no obvious duty to support. Few cases, however, are likely to be of this type. Just because we cannot help someone to recognise their own failure of recognition at one time does not mean that we cannot do so over a longer time period. Therefore, despite the entanglement of the paradox, the duty to support will still exist in most cases. It will, however, be limited in another way; for it cannot be used to justify decisions against a person's will. Before discussing this limit, however, it is worth evaluating whether the Mental Capacity Act recognises the duty at all.

The Mental Capacity Act may seem to recognise the duty to support. After all, section 1(3) requires 'all practicable steps' be taken to support someone to make a decision; and 4(4) that someone is encouraged and permitted to improve their ability to take part in decisions affecting them, insofar as that is 'reasonably practicable'. These provisions do not, however, require fulfilment of the duty to support; and this is, in part, due to the same 'decision-specific' structure of the Act that buttresses the duty of identification. Section 1(3) only requires that someone be helped to make the particular decision in question. It does not create any duty that continues after the decision is made; but it is then that the duty of support arises. Similarly, section 4(4) requires that the person's 'ability to participate' be supported during best interests decisions; but it, too, is specific to the particular decision

[72] Martha C Nussbaum, 'Aristotle, Nature, and Ethics' in JEJ Altham and Ross Harrison (eds) *World, Mind, and Ethics: Essays on the Ethical Philosophy of Bernard Williams* (Cambridge University Press 1995) 86, 100.

being made. It does not create any on-going duty.[73] Where a cohesive team is working well with the person, then an on-going duty to support might, nevertheless, emerge from these sections. As Series points out, however, both decision-making and support under the MCA are 'dispersed over a large number of disparate actors',[74] so the Act does nothing to encourage this sort of practice. Therefore, as with the duty of identification, the Mental Capacity Act allows the duty to support to be recognised, but does not require it.

It may be that the Mental Capacity Act is the wrong place to look for the duty to support. The provision of support, beyond that required for the immediate decision, could simply be beyond its remit. After all, as Lady Hale says, the Act is 'concerned with enabling the court to do for the patient what he could do for himself if of full capacity, but it goes no further'.[75] This point has been reinforced in the Court of Appeal: acting in someone's best interests grants no additional power 'to obtain resources or facilities from a third party'.[76] Given this, it is worth again looking to the Care Act 2014. After all, it gives local authorities duties to promote 'control by the individual over day-to-day life'[77] and to 'reduce the needs for care and support of adults in its area'.[78] At first sight, it seems possible to read a duty to help someone to recognise their own needs into these provisions; but, unfortunately, the detail of the Act makes this unlikely to consistently happen. Local authorities only have a duty to provide for needs that meet the 'eligibility criteria';[79] and these criteria refer only to first-order needs, such as maintaining nutrition.[80] Some service providers may decide that supporting someone to identify their own needs in the long term is the best way to help them to meet those needs; and local authorities can provide services when they have no

[73] Indeed, it only arises when 'the time for supported decision making is past'. *Re NRA* [2015] EWCOP 59, [2015] COPLR 690 [56] (Charles J).
[74] Series (n 5) 84.
[75] *Aintree University Hospitals NHS Foundation Trust* v *James* [2013] UKSC 67, [2013] 3 WLR 1299 [18].
[76] *Re MN* [2015] EWCA Civ 411, [2016] Fam 87 [80] (Munby P). Since this chapter was published as a paper the case has been approved in the Supreme Court: *Re N (An Adult) (Court of Protection: Jurisdiction)* [2017] UKSC 22, [2017] AC 549 [44].
[77] Care Act (n 49) s1(2)(d).
[78] *Ibid* s2(1)(c).
[79] *Ibid* s18(1).
[80] The Care and Support (Eligibility Criteria) Regulations 2014, SI 2015/313, reg 2(2) The regulations do mention 'accessing and engaging in work, training, education or volunteering' – reg 2(2)(h) – but this refers to someone's ability to access existing programs. There is nothing to indicate that it requires the provision of the support discussed here.

duty to do so.[81] Nevertheless, as with the MCA, there is nothing in the Care Act that makes fulfilling the duty to support required, instead of merely allowed.

Beyond failing to require the duty to support, the MCA also fails to recognise the limits of support. Words like 'support' can be slippery; and it is easy to envisage a duty to help someone to recognise that they have not recognised their needs quietly becoming a power to force someone to accept our assessment of their needs. Berlin's famous warning, 'to manipulate men, to propel them towards goals which you – the social reformer – see, but they may not, is to deny their human essence',[82] is as apt here as it is anywhere. One way to avoid this slippage is to return to the paradox, and to remember that deciding against a person's will is only justified when they do not recognise their failure to recognise a first-order need. If making a decision against someone's will does not make it possible to satisfy the first-order need, then this justification fails. This is what happens with the duty to support. If we say 'Due to the paradox, John cannot be helped to recognise his need for surgery, so we must decide for him', then a decision to operate can be implemented without any further input from John. In contrast, if we say 'John cannot be helped to recognise his needs, so we must decide for him', then we cannot make recognition happen without John's input, no matter what we decide. The subjective character of recognition dictates that it can only be reached by minimally consensual means; so the duty to support cannot justifiably be extended to decisions against a person's will. The MCA, however, does not distinguish between these two different situations; so it permits the meaning of 'support' to slip beyond reasonable bounds. For instance, in *Northamptonshire Healthcare NHS Foundation Trust* v *ML*,[83] a young man, 'ML', was given an 'opportunity to fulfil his potential' to be independent. The 'opportunity' involved a detention of up to two years in hospital and involuntary treatment that he would undoubtedly find traumatic.[84] It is unlikely that he experienced this as an opportunity or as support; and if he did not, then it must be doubted whether it really was these things. The MCA both fails to require respect for the duty of support, and fails to prevent 'support' sliding into unjustified coercion.

[81] Care Act (n 49) s19(1).
[82] Isaiah Berlin, 'Two Concept of Liberty' in *The Proper Study of Mankind: An Anthology of Essays* (first published 1958, Pimlico 1998) 191, 203.
[83] [2014] EWCOP 2, [2014] COPLR 439 [45] (Hayden J).
[84] *Ibid* [24], [59].

6. Conclusion

Meno's paradox clarifies one condition necessary for the justification of a best interest decision against someone's will. It suggests that it is not enough that someone has failed to recognise a first-order need, they must also fail to recognise their own failure of recognition. From this necessary condition, two duties can be derived: a duty to identify that any particular person is in this position before deciding against their will; and a duty to support those decided for to avoid this situation in the future. Any particular legislation that allows best interests decisions against a person's will can be measured against these duties. The Mental Capacity Act allows the fulfilment of these duties, but it does not robustly require them. Given this, it is perhaps unsurprising that the recent House of Lords report found the 'empowering ethos' of the Act's general principles has yet to be realised.[85] The failure to require compliance with the two duties is not due to those principles. It is, rather, due to the finer details of the Act and other legislation; some of which could, perhaps, be changed. The duty of identification requires that, during capacity assessments, the person's testimony is treated primarily as evidence of what they assert and only secondarily as evidence that they cannot recognise their needs. Changes to the legal framework governing the provision of advocacy has already brought the law closer to doing this; but there are important limits to what advocacy can achieve by itself. The second duty, the duty to support, could be given a legal basis by making a person's inability to recognise their other needs an eligible need under the Care Act. The MCA, however, also allows 'support' to slip into unjustified coercion, and this problem would have to be addressed separately.

This chapter shows that the Mental Capacity Act only allows, and does not robustly require, respect for two duties that any liberal system which allows best interests decisions against a person's will should be committed to. It does not, however, demonstrate whether or not best interests decisions are ultimately justified; and, given this, it does not support either the maintenance of the status quo, or the abandonment of all best interests decision making that has been called for under some interpretations on the UNCRPD. Nevertheless, Meno's paradox also presents a challenge to those who favour the latter view; for they would appear to be committed to one of three following responses to it: either that no-one is ever caught in the paradox; or that those who are so caught should be allowed to suffer whatever

[85] Select Committee (n 28) para 104.

avoidable harms follow from their second-order failures of recognition; or that third parties can interpret someone's 'will and preferences' in ways that contradict what the person actually says. A detailed examination of these positions is beyond the remit of this chapter; but the first response appears false, the second callous, and the third sinister. Despite this, taking seriously the response to the paradox discussed here would seem to narrow, without entirely closing, the gap between the MCA and the UNCRPD. It would require increased attention to the person's perspective during capacity assessments, and at least one form of additional, consensual support. These are both things that align well with the overall aims and ethos of the UN Convention.

Chapter Seven

Humility when responding to the abuse of adults with mental disabilities[1]

Published paper: (2017) 53 International Journal of Law and Psychiatry 102

1. Introduction

Legal theorists often reduce the ethics of responding to the abuse of another person to a clash between autonomy and protection.[2] This causes a problem, one of a type observed by Williams.[3] Theory aims for systemic unity, so it reduces the number of ethical ideas to a manageable minimum. 'Reflective criticism', in contrast, aims for shared understanding, so it uses any material that 'makes some sense and commands some loyalty'.[4] In other words, a choice must usually be made between theoretical simplicity and doing justice to the complexities of human life. This chapter is directed to the latter end. It suggests that responding to suspected abuse requires humility, something that cannot be usefully reduced to protection or respect for autonomy. It makes this argument in a series of widening concentric circles. Starting with a potential responder's attitude to their own knowledge, it expands to humility in individual actions, then to humility and legislative actions, and finally to humility and the UN Convention on the Rights of Persons with Disabilities ('the UNCRPD').[5] Throughout, the argument is supported with examples taken from the Court of Protection of England and Wales.

A lack of humility when responding to abuse can lead to systematic injustices for people with mental disabilities. It does not, however, do so in an obvious way. This is not a case of

[1] This paper forms part of a special edition on Article 16 of the United Nations Convention on the Rights of Persons with Disabilities. The full contents of this special edition may be found at http://www.journals.elsevier.com/international-journal-of-law-and-psychiatry/

[2] This binary is ubiquitous, but for examples see: James W Ellis, 'Decisions by and for People with Mental Retardation: Balancing Considerations of Autonomy and Protection' (1992) 37 Villanova Law Review 1779; Generva Richardson, 'Balancing Autonomy and Risk: A Failure of Nerve in England and Wales?' (2007) 30 International Journal of Law and Psychiatry 71; MC Dunn and others, 'To Empower or to Protect? Constructing the 'Vulnerable Adult' in English Law and Public Policy' (2008) 28 Legal Studies 234.

[3] Bernard Williams, *Ethics and the Limits of Philosophy* (first published 1985, Routledge 2011).

[4] *Ibid* 129-30.

[5] Convention on the Rights of Persons with Disabilities (adopted 13 December 2006, entered into force 3 May 2008) 2515 UNTS 3.

direct discrimination, or even of the discriminatory application of formally neutral laws, although both of those things also occur. Instead, it is a case where the background conditions of society mean that applying apparently neutral, but inadequate, systems in an apparently neutral way will nevertheless disproportionately affect a particular group. People with mental disabilities suffer more abuse than those without.[6] This is a primary systematic injustice. If responses to abuse, however, are inadequate, then this also leads to a secondary systematic injustice. Mentally disabled people, because they suffer disproportionate abuse, disproportionately rely on social and systemic responses to abuse. If those responses are inadequate, then the burden of that inadequacy will disproportionately fall on them. This secondary injustice could, of course, be addressed by stopping the abuse; but it could also be addressed by improving the flawed system. Given the probable difficulty of ending widespread abuse, it may be best to attempt both of these things.

Sections three and four show, by example, that the principles of autonomy and protection cannot adequately guide responses to abuse. The reason why they are inadequate can be easily summarised. Analysing a case of suspected abuse in terms of this binary focusses attention on the relationship between the potentially abused person and the potential abuser, but this oversimplifies the situation. For a response to be possible, then there will not only be a relationship between two people, abused and abuser; but a set of relationships between at least three: abused, abuser, and potential intervener. The particular features of a potential intervener, such as what they know and what they can do, cannot be ignored when deciding how they should respond to suspected abuse. If that is so, though, then any potential intervener should not only evaluate the situation in front of them. They should also evaluate their own relationship to that situation. They should, as Nagel says, 'step back from [their] initial view… and form a new conception which has that view and its relation to the world as its object'.[7] This 'step back' is not protection or respect for autonomy. It is something different, a form of self-evaluation. It is not enough, however, to merely say that potential responders to abuse should evaluate themselves. That would provide no information about which parts of the self should be evaluated or to what standards it should be held. If the idea of self-evaluation is to be useful, more detail must be provided. That

[6] Karen Hughes and others, 'Prevalence and Risk of Violence against Adults with Disabilities: A Systematic Review and Meta-Analysis of Observational Studies' (2012) 379 The Lancet 1621.
[7] Thomas Nagel, *The View from Nowhere* (Oxford University Press 1986) 4.

might seem like a massive task. Fortunately, though, there is no need to start from scratch. History provides a rich concept of self-evaluation that has been analysed for centuries: humility.

2. Humility

This section describes some features of humility.[8] Later ones use this description to analyse responses to abuse. One feature of humility, and enduring theme of the literature, is central to this chapter: humility includes an awareness of, and respect for, your own limits.[9] If humility is to augment the idea of self-evaluation, then this can provide a starting point. It suggests which part of themselves a potential responder to abuse should evaluate, their limits. Beyond this, however, humility can also help to provide the standards with which someone can evaluate themselves. This can be shown by analysing a succinct statement by Aquinas: 'humility restrains the appetite from aiming at great things against right reason'.[10]

Aquinas's statement has three important features. First, humility 'restrains the appetite', it leads someone to re-evaluate a course of action that they would otherwise be inclined to take. In the context of suspected abuse, this can work in either direction. Humility can moderate a desire to protect someone or a desire to respect their autonomy, when either is against 'right reason'. This is the second point. Humility does not restrain desire indiscriminately, it targets only desires that are unreasonable in the circumstances. Third, and finally, the ends that humility restrains someone from aiming at are 'great'. Humility is not important because it stops someone from trying to do wrong, other virtues do that, but because it stops them from unreasonably trying to do right. The desire to protect another person and the desire to respect their autonomy are both admirable. It is, however, possible to attempt either in an unreasonable way, as examples in later sections illustrate.

[8] Humility has a long and complex history. The aim here is to draw attention to some relevant details, not to give a complete account of the virtue.

[9] Compare three statements almost evenly spaced through the last seven centuries: Aquinas (1274) 'Knowledge of one's own deficiency belongs to humility'; Descartes (1649) 'We have humility as a virtue…as a result of reflecting on the infirmity of our nature'; and Grenberg (2005) 'The humble person understands her limits'. St Thomas Aquinas, *Summa Theologica* (first published 1274, Fathers of the English Dominican Province tr, Catholic Way Publishing 1947) IIa-IIae, 161, 2; René Descartes, *The Passions of the Soul* in *The Philosophical Writings of Descartes Volume 1* (first published 1649, John Cottingham and others tr, Cambridge University Press 1985) art 155; Jeanine Grenberg, *Kant and the Ethics of Humility* (Cambridge University Press 2005) 162.

[10] *Ibid* IIa-IIae, 161, 1.

Aquinas's statement makes it possible to address some persistent misunderstandings about humility. It is important to do so. Humility does not have the prominence that it had in earlier eras. Consequently, modern scholars can fail to grasp not only the subtlety of older traditions but also the central role of human limitation in enlightenment philosophies,[11] which are sometimes assumed to have consigned those traditions to history. This is not to say that humility is now entirely neglected, for it has been defended as essential to democracies,[12] especially pluralistic ones.[13] Nevertheless, three misunderstandings are common: a conflation of humility and self-abasement; the idea that humility necessarily reinforces existing hierarchies; and the idea that humility is overly inwards and self-directed.

The first common misunderstanding is to equate humility with self-abasement, so to think that it must necessarily lead to inaction or despair. This view fails to notice that someone is only humble when they recognise a limit that they actually have. Being aware of real limits does not require someone to imagine limits that do not exist. When Murdoch says that 'humility is not a peculiar habit of self-effacement, rather like having an inaudible voice, it is selfless respect for reality',[14] her point is orthodox. The emphasis on right reason in Aquinas has already been mentioned; and, similarly, Augustine states that humility should be 'on the side of truth, not on the side of falsity'.[15] Tying humility to honesty in this way undermines the idea that humility leads to inaction, and this is significant for responses to abuse, where anything that leads to unnecessary inaction might seem suspect. The connection to honesty also reveals something of the deeper rationale for humility. One reason that humility is important is that humans do systematically tend to overlook their own limits. This insight is old, but experimental psychology tends to support it;[16] and just as importantly, psychology also suggests that this overconfidence can sometimes be addressed.[17] In other words, science seems to reaffirm both the need for humility and its

[11] Jennifer Clement, *Reading Humility in Early Modern England* (Routledge 2015) 15-20; Julie E Cooper, *Secular Powers: Humility in Modern Political Thought* (University of Chicago Press 2013) 5.
[12] Mark Button, '"A Monkish Kind of Virtue?" For and Against Humility' (2005) 33 Political Theory 840.
[13] Erik Parens, 'The Pluralist Constellation' (1995) 4 Cambridge Quarterly of Health Care Ethics 197.
[14] Iris Murdoch, *The Sovereignty of Good* (first published 1970, Routledge 2014) 93.
[15] St Augustine, *Nature and Grace* in JE Rotelle (ed), *The Works of Saint Augustine Volume 23* (first published 415, RJ Teske tr, New City Press 1997) 197, ch 38.
[16] David Dunning, *Self-Insight: Roadblocks and Detours on the Path to Knowing Thyself* (Taylor and Francis 2005).
[17] Joyce Ehrlinger and others, 'Understanding Overconfidence: Theories of Intelligence, Preferential Attention, and Distorted Self-Assessment' (2016) 63 Journal of Experimental Social Psychology 94.

possibility. This, too, is relevant to responses to abuse, where demanding the impossible would be little practical help.

The second common misunderstanding of humility is to think that it necessarily maintains hierarchies; that it is something forced on the weak by the strong. This misunderstanding, like the last, is deeply ahistorical. It does not pay attention to where the advice to be humble was traditionally directed. Although the virtue has been invoked to support existing social orders, there are also influential traditions that assert it is the powerful who most need humility if their power is to remain legitimate. Clement shows that humility in the early modern period was widely understood as 'a virtue that resists tyranny even as it can be invoked to support the status quo';[18] and, similarly, Klancer shows that for both Aquinas and the Confucian Zhu Xi, humility is 'a virtue inherently fitting for the strong'.[19] The idea that it is the powerful who must be humble has more recently been applied to clinical relationships;[20] and this aspect of the virtue is important to the present chapter, for those in a position to intervene in cases of abuse almost invariably have considerable power relative to the person thought to be suffering abuse.

The final common misunderstanding of humility is to think that because it is concerned with the person's own limits that it is too inwardly directed to usefully guide action. Once again, this idea lacks historic support. As Button notes,[21] even St Bernard, a monk writing primarily for other monks, emphasised humility not as an end in itself, but as a precondition to cultivating good relationships. Indeed, Bernard was blunt: 'you will never have real mercy for the failings of another until you know and realise that you have the same failings in your soul'.[22] Similarly, Boyd points out that for Aquinas, too, humility is social. It 'enables us to value ourselves as members of a community in which no one person possesses independent, god-like status'.[23] The more recent literature has, if anything, further emphasised the social implications of humility: Button argues that pluralistic democracies depend on citizens cultivating sensitivity to their own limits,[24] and the limits

[18] (n 11) 127.
[19] Catherine Hudak Klancer, 'How Opposites (Should) Attract: Humility as a Virtue for the Strong' (2012) 22 The Heythrop Journal 662, 670.
[20] Sayantani DasGupta, 'Narrative Humility' (2008) 371 The Lancet 980.
[21] (n 12) 850.
[22] Bernard of Clairvaux, *The Steps of Humility and Pride* (first published 1120, M Ambrose Conway tr, Cistercian Publications 1973) III.6.
[23] Craig A Boyd, 'Pride and Humility: Tempering the Desire for Excellence' in Kevin Timpe and Craig A Boyd (eds), *Virtues and their Vices* (Oxford University Press 2014) 245, 258.
[24] (n 12).

of their institutions and laws; Penrose that humility underwrites the correct moral response to wrongdoers;[25] and Coulehan that humility is necessary to respond appropriately to the 'ambiguities, mysteries, and surprises' of the clinical encounter.[26] This chapter is a further small contribution to the social argument for humility. Its central claim is that only humble responses to abuse will be socially effective, both at the level of the abused person's relationships and at the level of the whole society. To show this, though, more detail is needed on the particular limits that a potential responder to abuse should recognise and respect. The next section introduces the first type of limit, limits to knowledge.

3. Humility about knowledge when responding to abuse

3.1 Humility about what is known

This is the first of two sections discussing humility when confronting a particular case of suspected abuse. It addresses humility about knowledge, or 'epistemic humility'. The need for this sort of humility is illustrated using a case from the Court of Protection of England and Wales, *Somerset County Council v MK*.[27] This case concerned a young woman, 'P', with severe learning disabilities and autism who had lived with her family for her entire life.[28] Bruising was noticed on P's chest, and this quickly led to an investigation by social workers that concluded that it was 'highly likely that P has received a significant injury from someone or something other than herself'.[29] Consequently, she was removed from her family and transferred to a 'respite placement' and then to an 'assessment and treatment unit'.[30] Clearly, this action was intended to protect P, but it could also be plausibly presented as furthering her autonomy. Abusive relationships do not respect autonomy; and, beyond that, 'autonomy-impairing fear' may be a major reason that such relationships continue.[31] If so, it can be plausibly argued that removing someone from an abusive relationship might promote their autonomy. In *Somerset*, then, protection and respect for autonomy could be easily made to work together to justify removing P from the family

[25] Brian Penrose, 'Humility and Understanding' (2010) 39 Philosophical Papers 427.
[26] Jack Coulehan, 'A Gentle and Humane Temper: Humility in Medicine' (2011) 54 Perspectives in Biology and Medicine 206.
[27] [2014] EWCOP B25.
[28] *Ibid* [3].
[29] *Ibid* [6] (Judge Nicholas Marston).
[30] *Ibid* [8], [11].
[31] Jennifer Nedelsky, *Law's Relations: A Relational Theory of Self, Autonomy, and Law* (Oxford University Press 2012) 312.

home. This is why epistemic humility was also needed. Just prior to the bruising being found, P had been witnessed hitting herself in the chest on a school trip and had been physically restrained by staff. These events would have been 'easily discoverable by the social workers if they had carried out a proper investigation'.[32] In the year that elapsed before these easily discoverable facts were recognized, P spent time deprived of her liberty in a placement that was 'clearly not appropriate'. This led to anti-psychotics being prescribed to manage her behaviour.[33]

The initial suspicion that P was suffering abuse was reasonable, but something clearly went wrong in the response to that suspicion. This wrongness cannot be adequately caught using only the concepts of protection and autonomy. The inadequacy of these two concepts, however, may not be obvious. It might be objected that removing P from her family did not protect her, it exposed her to new risks of emotional harm and inappropriate medication. Similarly, it did not respect her autonomy, it required overriding her consistently expressed wish to go home.[34] The social workers, according to this argument, failed to protect P or to respect her autonomy; so these concepts are sufficient to express what went wrong. This analysis, however, rests on a fundamental error. Criticizing the social workers for failing to protect P or respect her autonomy assumes that they knew then what we know now. To assume this is, however, to miss the point in the most dramatic way possible. The social worker's culpable fault was that they *did not know* something that they should have known. If it was possible to travel back in time and talk to the professionals who interacted with P, then it would not help to tell them to protect her. All of the evidence indicates that they thought that they *were* protecting her. Similarly, it would do no good to tell them to respect her autonomy. The evidence also indicates that they thought that they *were* respecting her autonomy. Indeed, even after conceding that there were no 'safeguarding issues', the local authority continued to argue that P should not be returned to her family, so that she could better fulfill 'her wider potential'.[35] This, however misplaced, is an argument from autonomy. Just because hindsight shows that the local authority failed to protect P or to respect her autonomy does not mean that these concepts could have adequately guided decision-makers at the time.

[32] (n 27) [6], [76].
[33] *Ibid* [10].
[34] *Ibid* [41].
[35] *Ibid* [45].

If it was possible to travel back in time and talk to the professionals who interacted with P, then it might have helped to draw their attention to how little they really knew about the situation. Had the social workers paid more attention to the limits of their knowledge, then it is reasonable to believe that they would have investigated more fully. In addition to evaluating how to protect P and respect her autonomy, they also needed to evaluate themselves. Indeed, in court, the senior social worker admitted this, saying that there was 'a failure of those on the ground to realize that they are out of their depth'.[36] Furthermore, there is evidence that the wider 'systemic failure'[37] of the local authority was also, at least in part, a failure of epistemic humility. Actions were repeatedly taken without consulting those who knew P best,[38] and P's own attempts to communicate her wish to go home were perversely misinterpreted.[39] Indeed, the way that the local authority was eventually criticized in court parallels Aquinas on humility. Judge Nicholas Marston and the senior social worker were both careful not to question 'the motivation of LA to do the right thing'.[40] Instead, their criticism focused on the 'misguided'[41] actions that were carried out towards that admirable purpose. In Aquinas's terms, the social workers aimed at a 'great thing'. Unfortunately, they did so 'against right reason'.

3.2 Humility about what can be known

Somerset demonstrates the need for humility about what is known, but epistemic humility has another aspect, humility about what it is possible to know in the circumstances. Another case, *A Local Authority v M*,[42] provides an example. The facts in this case are 'long and complex',[43] but a bare outline will suffice here. 'M' is a young man with a learning disability and autism. He lived at home until he was 18. After that, he was moved between his home and various residential placements.[44] The case concerned a series of allegations and counter-allegations between his parents and the local authority. The authority alleged that M's mother was excessively controlling, to an extent that caused M harm;[45] M's parents counter-alleged that the managers of the residential placements were both controlling and

[36] *Ibid* [74].
[37] *Ibid* [83].
[38] *Ibid* [8],[10],[12].
[39] *Ibid* [40]-[41].
[40] *Ibid* [58].
[41] *Ibid.*
[42] [2014] EWCOP 33, [2015] COPLR 6.
[43] *Ibid* [3] (Baker J).
[44] *Ibid* [1], [24], [31], [36], [39], [53].
[45] *Ibid* [2], [24], [60], [191].

neglectful.[46] Although Mr Justice Baker found overwhelmingly for the authority,[47] he made it clear that the case was evidentially complex,[48] and one evidential difficulty is relevant here. In the words of Dr Carpenter, a consultant psychiatrist:

> 'At present, on any choice [M] makes, one is not totally confident that what is chosen is what he truly wants as his passive acceptance of most instructions makes it possible that, even if he did not really want what he chose, his continuation to use or eat or perform that choice could be evidence of his compliance as much as true desire'.[49]

In other words, attempts to determine M's non-coerced choices may have been experienced by M as further coercion. Here, a difficulty emerges for any analysis based narrowly on respect for autonomy and protection. The local authority and M's parents counter-alleged 'controlling behaviour'; so, as in *Somerset*, protection and respect for autonomy did not necessarily clash. Interferences with M's autonomy were what he was to be protected from. To protect M's autonomy, however, required some idea of when M was autonomous; and this could not be determined. Furthermore, what thwarted the investigation was the same thing as prompted it, the fear that a controlling environment led to M merely complying with authority figures.

Unlike *Somerset*, *M* does not just show a limit to what a potential intervener knew. It shows a limit to what they could know in the circumstances. This requires a different sort of response. A limit to what the intervener knows should usually prompt further investigations. In contrast, a limit to what can be known should prompt the abandonment of that line of investigation, the opening of other lines of enquiry, and great caution when acting. Even here, though, it is important not to let humility be confused with despairing inaction. Limits to knowledge in one respect and at one time need not be limits in every respect or forever. In *M*, Baker J endorsed the need for ongoing support to help M to articulate his own wishes.[50] In time, that might escape the problem that Dr Carpenter identified. In the meantime, the judge also examined other evidence to determine which environments were likely to be unduly controlling. Both of these moves illustrate an

[46] *Ibid* [31], [63], [68], [205].
[47] *Ibid* [211]-[219].
[48] *Ibid* [72]-[76].
[49] *Ibid* [138].
[50] *Ibid* [138].

important point. Even where epistemic humility leads to a recognition that something cannot be known, it need not lead to inaction. Instead, it can prompt better responses to suspected abuse. This is also true of another type of humility, humility about action.

4. Humility and action when responding to abuse

4.1 Humility about action

This section, like the last, examines humility when responding to a particular case of suspected abuse. This subsection addresses humility about action, respect for the limits to what an action can achieve. As with humility about knowledge, humility about action can guide a potential responder to abuse in ways that attention to autonomy and protection alone will not. The case of *The London Borough of Tower Hamlets v TB*[51] illustrates this point. It concerned a 41 year old woman with a learning disability,[52] whose husband physically and verbally abused her.[53] Although the judgement also addresses TB's capacity to consent to sex and contact with her husband, and the question of depriving her of her liberty, it is the court's treatment of her residence that shows the need for humility about action. The reported case occurred about two years after TB had been placed in supported accommodation provided by the local authority, but it quotes at length an earlier interim judgment authorising that placement. This juxtaposition of two lines of reasoning separated by two years provides something very rare: a case in which we know both why a decision was reached in the Court of Protection and, in some detail, what happened next.

It is plain that a need to protect TB from 'the risk of physical violence' partially motivated moving her into supported accommodation.[54] It was not, however, the only factor in that decision. In the earlier hearing, TB was described as living an 'empty life' in 'a wholly oppressive environment where there is a sense of containment', which 'revolved only about …watching television'.[55] Furthermore, there 'were no skills being developed'.[56] These observations formed the basis of an argument from autonomy; emphasising the oppressiveness of TB's current life laid the foundations for an argument that moving her would enhance her ability to be self-determining. Indeed, Mr Justice Mostyn was careful to

[51] [2014] EWCOP 53, [2015] 2 FCR 264.
[52] *Ibid* [1]-[2].
[53] *Ibid* [5(9)].
[54] *Ibid* [5], [23].
[55] *Ibid* [5(13)].
[56] *Ibid.*

disarm any countervailing argument based on TB's occasionally expressed desire to live in the marital home, saying 'the wishes and feelings of TB have varied and seem to be the product of the environment in which she is presently sited'.[57] On this reading, such wishes were not genuinely autonomous. In the interim judgment, then, it was found that a move into supported living would not only protect TB from her abusive husband but also enhance her autonomy. As in *Somerset* and *M*, autonomy and protection were made to harmonise.

In *Somerset*, protecting P and valuing her autonomy both seemed to require removing her from her home, but epistemic humility could have prompted a proper investigation of the cause of her bruising. Similarly in *TB*, both principles seemed to require removing TB from her home, but humility could have prompted a more sophisticated response. In this case, however, it was the limits of action, not knowledge, which did not receive attention. This is because a common asymmetry occurred in analysis. It was observed in some detail that TB's living situation did not afford her much autonomy. There was, however, no real analysis of whether the proposed action, moving her, was likely to increase her self-determination. Consequently, the limits to what could be achieved just by removing TB from a site of overt abuse were not taken into account. This led to professionals underestimating what a commitment to increasing TB's autonomy would entail. Two years after the initial decision, Mr Justice Mostyn observed that she had 'not prospered there as much as everyone hoped. She seems to lead a rather isolated and lonely life, spending hours lying on the sofa watching TV'.[58] At that point, the need to provide her with additional support was belatedly recognised.[59]

TB does not provide as striking an example of moral error as *Somerset*, although two years of inadequate support should not be discounted. TB was, at least, protected from further domestic violence. It is, however, worth examining how she was protected. The focus on her lack of autonomy allowed the issue of protection to become distorted in a familiar way. As has been observed in other cases, it made the central problem not the abuser's actions but a purported 'defect' in the abused person's 'abilities to resist external pressure'.[60] The

[57] *Ibid* [5(14)].
[58] *Ibid* [8].
[59] *Ibid.*
[60] Laura Pritchard-Jones, 'The Good, the Bad, and the "Vulnerable Older Adult"' (2016) 58 Journal of Social Welfare and Family Law 51, 60.

best way to respond to TB's husband's violence was therefore seen to be removing *her* from the house and depriving *her* of her liberty. Without rhetoric about promoting TB's autonomy this action looks, rightly, strange. This rhetoric, though, could be nothing more than hyperbole without humility about action. Only paying attention to the limits of any particular actions would have led to professional's recognising what TB needed to be self-determining. Once again, this criticism echoes Aquinas. To aim at increasing TB's autonomy was a 'great thing', but to think it could be so easily achieved was against 'right reason'. This example, too, deflates the stereotype of humility leading to inaction. Instead, it may have led to better action. Two years later, when it was recognised that TB needed more support, it probably did.

4.2 Humble actions

Humility about action is not the only connection between humility and action. There is also a need for humble actions. In other words, it is not enough for someone confronting suspected abuse to reflect inwardly on limits, whether of their knowledge or of their possible actions. They must also acknowledge those limitations to the people that they might affect. In the context of responses to suspected abuse, this will especially be the person thought to be at risk of abuse. The idea of humble actions is nothing new. Aquinas recognises the value of 'outward acts' of humility performed with 'due moderation'.[61] Three further points are, however, significant in the current context: humble actions can sometimes help to overcome the limits that they acknowledge; they raise the danger of false humility; and there is reason to suspect an endemic lack of this public humility in some services in England and Wales.

Humble actions can sometimes help to overcome the limits that they acknowledge. Although not directly concerned with abuse, two passages by judges in the Court of Protection exemplify this point. Each concerns the effect on the judge of going to hospital to meet a person that they might be making decisions for, and each has something of the character of a personal testimony. After visiting the person, Mr Justice Peter Jackson said 'I obtained a deeper understanding of Mr B's personality and view of the world';[62] and, similarly, Mr Justice Mostyn 'the person I met was different in many respects to the person

[61] (n 9) IIa-IIae, 161, 3.
[62] *Wye Valley NHS Trust v B* [2015] EWCOP 60, [2015] COPLR 843 [18].

described in the papers'.[63] It is easy to misunderstand what has happened here. Because the person in question is being allowed to express a preference it might seem that the judge's actions can be completely understood by reference to autonomy. This is not so. Going to meet someone probably does help a judge to better respect that person's autonomy, but it does so because it is an expression of judge's awareness of their own limits. This expressive aspect of humility is important. When a High Court judge, with the real and symbolic power inherent in the role, leaves their courtroom to listen to someone with a mental disability, with all the stigma and history of that diagnosis, then an important reversal has taken place. Furthermore, by publically acknowledging the limits to their knowledge of the person, these judges have allowed themselves to partially address that limitation.

Humility is necessary to address the limits that it acknowledges, but it is seldom enough. For meeting the person to make any real difference, the judges needed a bundle of further skills and virtues. For instance, they had to be approachable enough able to put the person at ease, attentive enough to hear their story, and open-minded enough to let their own assumptions be challenged. Still, humility is critical in two ways. First, without the initial act of expressive humility, the judge simply would not have had the opportunity to exercise the other skills. Second, many of these skills require some humility to develop. To be good at listening to others requires practice, but only those that have acknowledged they could be better at listening to others will actively engage with this practice. Although humility is not usually enough to address the limits it acknowledges, it can help create both the opportunity and the skills necessary to address those limits.

If humble actions can help to overcome the very limits that they acknowledge, then they can be surprisingly potent. This potency can, however, lead to dangers. If public displays of humility can be powerful, then there is a potential motive for false humility. The *Somerset* case provides a striking example:

> …the family were invited to discuss plans about P's future and express their views. In fact it is clear that was not the reason they were invited at all. Far from a change of heart and an attempt to communicate the reason is clear. It was felt by Mr M on advice from the LA lawyers that: "The COP might pick up that no

[63] *A Hospital NHS Trust v CD* [2015] EWCOP 74, [2016] COPLR 1 [31].

'round table' meeting has been held and this might disadvantage us during the hearing"[64]

Here, something that seemed to show that the local authority valued the families input, and therefore seemed to be based on some recognition that the 'experts' did not have all the answers, was a deliberate deception. The humility that the local authority was displaying was false in at least two ways. First, it was false in the direct sense that the professionals only seemed to acknowledge their limits. There was no genuine interest in learning from the family, just a desire to win a court case. Second, it was false in terms of audience. The act of holding a meeting appeared to address the family, but it was really addressed to the Court of Protection. False humility may be an endemic danger in social contexts where humility is valued. For example, in early modern England, where humility was highly prized, false humility seems to have been considered an almost ubiquitous threat.[65] Nevertheless, the danger of false humility should not be overstated. Judge Nicholas Marston was not deceived by the authority's false humility in *Somerset*, and it seems unlikely that the family were either. Indeed, far from false humility helping the professionals to get what they wanted, it eventually helped to condemn them.

There may be a lack of sufficiently humble actions in some services. This chapter is not the place for any systematic study of this phenomenon. All the same, it is easy to find a worrying example: professional behaviour around the premature deaths of people with learning disabilities. People with learning disabilities are more likely than others to die from causes that could be avoided with good-quality healthcare,[66] and a lack of knowledge of the individual is often the critical factor in this process[67]. Furthermore, the investigations needed for effective diagnosis are often neglected; and the concerns of people with learning disabilities, their families, and carers, are not taken seriously.[68] This suggests a lack of epistemic humility, a failure by professionals to acknowledge the limits of their knowledge. It also, however, suggests a failure to act humbly. Learning from someone with a learning disability, or from their family, often requires a professional to admit that in some areas

[64] (n 27) [70].
[65] Clement (n 11) ch 2.
[66] Pauline Heslop and others, The Confidential Inquiry into Premature Deaths of People with Learning Disabilities (CIPOLD) (Norah Fry Research Centre 2013).
[67] *Ibid* 83.
[68] *Ibid* 57. Indeed, there is evidence that this failure to listen is an even more widespread problem: Peter Beresford and others, *From Mental Illness to a Social Model of Madness and Distress* (Shaping our Lives 2016).

these people have more expertise than they do. A failure to take their concerns seriously does not only reflect an inward failure to recognise limits. It is also a public failure to act with appropriate humility. Furthermore, this failure to act humbly even seems to persist after someone with learning disabilities has died prematurely. A review of one NHS Foundation Trust found that despite people with learning disabilities in contact with it having an extremely low average age of death, it systematically failed to investigate their unexpected deaths.[69] Here, again, a failure of epistemic humility was intimately tied to a failure to act humbly; for instance, by listening to the concerns of families.[70]

Reports into the deaths of people with learning disabilities raise extremely hard questions for professionals. People with learning disabilities consistently die young, and their early deaths have been consistently tied to professional failings.[71] Furthermore, the implicated patterns of professional behaviour continue even after someone with learning disabilities has died.[72] In these circumstances, it does not seem overly dramatic to ask whether, far from protecting people with learning disabilities from abuse, the state is systematically perpetrating abuse. If so, much of this abuse may be inflicted unwittingly. The professionals involved may, in a pattern familiar to social psychologists,[73] not only fail to recognise their own limits but also fail to recognise that they have done so. This is another problem that the principles of autonomy or protection cannot solve alone. Before someone dies, a professional without humility can wrongly assume that they are that person's protector or that they empower the person to be autonomous; and, after someone dies, the same professional might conclude that there is now no-one that they need to protect or empower. Humility, in contrast, directs attention to the common source of the problem, the professional's unacknowledged limits.[74] In this way, once again, humility can help to shape interpersonal relationships: not only how someone directly responds to abuse but also potentially abusive institutional relationships. This ability to shape interpersonal relationships has implications for the more indirect relationships of the law.

---

[69] Bob Green and others, *Independent Review of Deaths of People with a Learning Disability or Mental Health Problem in Contact with Southern Health NHS Foundation Trust April 2011 to March 2015* (Mazars 2015) 16-17.
[70] *Ibid* 32.
[71] Heslop (n 66).
[72] Green (n 69).
[73] Dunning (n 16) ch 2.
[74] A sufficiently arrogant professional might nevertheless think that they are excellent at humility. Hopefully, that thought is contradictory enough to worry all but the most narcissistic individuals.

5. Humility and the law

5.1 Humility and disability discrimination

The previous two sections show the need for humility when directly responding to abuse. This section explores the implications for law. First, this subsection shows that a lack of humility in responses to abuse will disproportionately affect people with mental disabilities, then later subsections examine the law's ability to foster humility. The discrimination point is important. The analysis so far only shows that a lack of humility when responding to abuse can be harmful, not that it will disproportionately affect those with disabilities. To see that it will, two complicating factors must be put to one side. First, the role of intention: in none of the cases discussed so far is there evidence that the professionals involved intended to discriminate against the person. Analysis in terms of humility will even concede this point. After all, saying that someone lacked humility admits that they aimed at a 'great thing'. A lack of discriminatory intent, however, does not mean that discrimination did not happen. 'Discrimination' in the UNCRPD includes discriminatory purpose *or* effect,[75] so unintentional discrimination remains discrimination.

The second factor complicating the link between humility and discrimination is the effect of domestic law, in this case the Mental Capacity Act 2005 ('the MCA'). The examples in this chapter were decided under the MCA, but Section 2(1) of the Act limits its application to those who are unable to decide *because* of 'an impairment of, or a disturbance in the functioning of, the mind or brain'. It is probably impossible to construe this in a disability-neutral way, and this context may give the impression that any discrimination must be due to this law, and not due to a lack of humility. Such an impression would be misleading. Merely changing the law would not be enough to avoid substantive discrimination. If the relevant section of the MCA was removed, then the 'functional' test in section 3(1) of the Act, which requires that a person 'understand' and be able to 'use and weigh' the information relevant to a decision, would remain. Under this section, a higher proportion of individuals with mental disabilities than those without would be found to lack capacity.[76]

[75] Art 2.

[76] Eilionóir Flynn and Anna Arstein-Kerslake, 'The Support Model of Legal Capacity: Fact, Fiction, or Fantasy?' (2014) 32 Berkeley Journal of International Law 124.

If that is so, though, then persons with disabilities would still disproportionately suffer the effects of insufficiently humble responses to suspected abuse.

Even if all perceived differences in understanding between people with and without mental disabilities are only an expression of stigma, then removing section 2(1) of the Act would not be enough to avoid discrimination. Any judgment about someone's ability to understand, or use and weigh, information is necessarily influenced by the social context in which it is made;[77] and widespread, strong, and unacknowledged negative attitudes towards people with disabilities seem to be part of that context.[78] Stigma cannot simply be legislated away. A radical response to this problem might be to suggest the removal of all powers to intervene in cases of abuse. This, however, would achieve formal non-discrimination at the cost of substantive equality. People with mental disabilities suffer more abuse than those without,[79] so a wholesale refusal to intervene would disproportionately affect them.[80] A more prudent response would be to suggest that that powers to intervene should be formally disability-neutral, effective, and carefully scrutinised. If, however, non-discrimination requires an effective general response to abuse, and an effective response to abuse requires humility, then legislative scrutiny must also pay attention to the need to foster humility.

5.2 Fostering humility with the law

At first, the idea of using legislation to foster humility can appear counter-intuitive. It is hard to imagine a law stating 'be humble', much less a judge ruling on whether or not a particular professional has complied with it. An important distinction must, however, be made: between calling for direct state enforcement of a virtue; and paying attention to the detail of how the law already promotes some virtues at the expense of others.[81] If humility is important, then it would be negligent not to do the latter. After all, the law already does

[77] Natalie F Banner, 'Can Procedural and Substantive Elements of Decision-making be Reconciled in Assessments of Mental Capacity?' (2013) 9 International Journal of Law in Context 71.

[78] Michelle Clare Wilson and Katrina Scior, 'Attitudes towards Individuals with Disabilities as Measured by the Implicit Association Test: A Literature Review' (2014) 35 Research in Developmental Disabilities 294.The science underlying the measurement of both bias and discriminatory outcomes has serious problems with both validity and reliability. Nevertheless, absence of proof is not proof of absence, and the existence of widespread disability discrimination remains, unfortunately, very plausible. Rickard Carlsson and Jens Agerström, 'A Closer Look at the Discrimination Outcomes in the IAT Literature' (2016) 57 Scandinavian Journal of Psychology 278.

[79] Hughes (n 6).

[80] Camillia Kong, 'The Convention for the Rights of Persons with Disabilities and Article 12: Prospective Feminist Lessons against the "Will and Preferences" Paradigm' (2015) 4 Laws 709.

[81] Nedelsky (n 31) 69-73. She uses the language of 'values' not 'virtues'. The point is made here by analogy.

influence professional responses to suspected abuse. If they sometimes seem insufficiently humble, then there is reason to ask whether the law sometimes fosters professional arrogance instead of humility.

The purpose of this chapter is to make the foundational point that humility is important when confronting suspected abuse, so it is not the place for a systematic analysis of the law of any jurisdiction. Nevertheless, the MCA provides a convenient illustration of the way that law can foster arrogance instead of humility, as the examples in earlier sections concerned its application. As is often the way with law, what the MCA omits is at least as significant as what it includes. Under the Act, a person's 'wishes and feelings' must be taken into account after it has been determined that they lack capacity to make a decision,[82] and the influence of the UNCRPD has caused this part of the Act to receive considerable attention.[83] The person's own perspective *during* a determination of incapacity has, however, received less recent attention. The Act does not allow incapacity to be assumed, inferred merely because of an unwise decision, or established merely by reference to someone's condition, behaviour, age, or appearance.[84] Inability to make a decision is, however, clearly taken to be an objective fact,[85] and the Act does not require any attention be paid to the person's own perspective on their capacity when determining this objective fact.[86]

If humility is important, then inattention to the person's perspective during capacity assessments is a problem. There are often significant limitations to how well someone can assess another person's understanding, especially on the short time-scale of many professional relationships. Indeed, sometimes the person being assessed might have a better grasp of their own understanding of a situation than the assessor does. Despite this, the law does not require that if someone disagrees with a capacity assessment, this be taken as evidence that the assessment might be wrong. Indeed, the opposite can happen, and disagreement can be taken as evidence that the person *lacks* capacity.[87] In these situations, capacity assessments can be experienced by the assessed person as authority figures simply

[82] MCA s4(6)(a).
[83] Wayne Martin and others, *Three Jurisdictions Report: Towards Compliance with CRPD Art 12 in Capacity/Incapacity Legislation across the UK* (University of Essex 2016) 39-42.
[84] MCA ss1(2), 1(4), 2(3).
[85] MCA s2(1).
[86] Paul Skowron, 'The Implications of Meno's Paradox for the Mental Capacity Act 2005' (2016) 24 Medical Law Review 279.
[87] Neil Allen, 'Is Capacity "In Sight"?'(2009) 19 Journal of Mental Health Law 165.

exerting power over them.[88] When the assessed person may have been abused, then this is especially worrying, the repetition of a pattern that they may know all too well. Of course, there is no reason to think that this is what always, or even usually, occurs. Many capacity assessors are acutely aware of their limits, and correspondingly attentive to the person's perspective. That, though, is despite the law, not because of it.

5.3 Legislative humility

The example of capacity assessments suggests that sometimes the law fails to foster humility, but it does not suggest that this process is inevitable. For example, the MCA could have required the person's own perspective be taken into account during assessment or included robust supports to contest a finding of incapacity. All the same, there are reasons to be cautious when suggesting legal reform. Passing legislation is an action; and, like any action, there are limits to what it can achieve. Legislative humility, respect for the limits of what law can achieve, is therefore needed. Indeed, some limits are intrinsic to the law. As Aristotle notes, 'all law is universal, and yet there are some things about which it is not possible to make correct universal pronouncements'.[89] In other words, no matter how certain and just a law may appear on paper, it will sometimes be uncertain or unjust when applied to a particular sets of facts. Sir Edmund Thomas may be right when he laments that this 'continues to elude so many in the legal community', whether practitioners or academics.[90] If he is right, then there is a profound danger here. Those unaware of their limits characteristically also fail to realise that they are unaware of their limits,[91] so an unrealistic legal community may be largely unaware that it is being unrealistic. A lack of legislative humility risks the position, in Taylor's words, where 'we can't see any more the way these rules fit badly our world of enfleshed human beings, we fail to notice the dilemmas they have to sweep under the carpet'.[92]

Taylor's warning is perceptive, but it is to be hoped that legal debate has not yet become so apocalyptically unrealistic. All the same, the need for legislative humility is real, as two

[88] Susan Stefan, 'Silencing the Different Voice: Competence, Feminist Theory and Law' (1993) 47 University of Miami Law Review 763, 782-5.
[89] Aristotle, *Nicomachean Ethics* (first published approximately 322BC, Christopher Rowe tr, Oxford University Press 2002) 1137b13-15.
[90] EW Thomas, *The Judicial Process: Realism, Pragmatism, Practical Reasoning and Principles* (Cambridge University Press 2005) 116-7.
[91] Dunning (n 16) ch 2.
[92] Charles Taylor, *A Secular Age* (Harvard University Press 2007) 742.

quotations about the MCA, separated by almost a decade, illustrate. In 2005, Lord Falconer, moving for a second reading of the Bill that would become the Act, argued that it was necessary because 'the current law is confusing, often incomplete, and much misunderstood'.[93] The Bill duly became law. Nevertheless, after the Act had been in force for seven years, a House of Lords Select Committee tasked with reviewing it concluded that it had 'suffered from a lack of awareness and a lack of understanding';[94] and observed widespread confusion and incompleteness.[95] The MCA, it seems, did not even effectively clarify the law, much less achieve any more ambitious aim.

For the most part, the Select Committee blamed 'implementation' or 'culture', not the Act itself, for its problems. To acknowledge problems with implementation is, however, to implicitly acknowledge the limits of law. It directs attention to the other things, beyond the law, that the Act requires to have effect. Even so, the Select Committee underestimated the limits of legislation. It generally assumed, when talking about implementation, that the MCA had established a clear rule that the 'culture' should be brought into alignment with.[96] This is too optimistic. As Sunstein says, 'the content and nature of a legal provision cannot be read off the provision'.[97] Interpreting a law requires social context. This, however, suggests that the idea that culture can be changed by bringing it into alignment with a rule is too simple, for the rule depends on its cultural context for its own meaning.[98] Law can affect culture, but it never does *this* by simply establishing a new rule; for the rule, if it is understood at all, is understood in the terms of the culture in which it is implemented. Instead of changing culture by setting standards, the law is, at most, one part of a slow and unpredictable conversation. This does not mean that legal responses to abuse cannot be improved. Legislative humility, like other forms of humility, does not prescribe inaction. It merely counsels scepticism about attempts to use the law to directly impose a new vision of society. This counsel is hardly new,[99] but it does require a particular sort of ideal. That ideal is not the creation, then implementation, of a blueprint for the prevention of all abuse.

[93] HL Deb 10 January 2005, vol 668, col 13.
[94] Select Committee on the Mental Capacity Act 2005, *Mental Capacity Act 2005: Post-Legislative Scrutiny* (2013-14, HL Paper 139) 6.
[95] *Ibid* paras 63, 97, 167, 267, 276; in 'Conclusions and Recommendations' paras 31, 46, 48, 70.
[96] *Ibid* paras 15, 38, 39, 84, 85.
[97] Cass R Sunstein, 'Problems with Rules' (1995) 83 California Law Review 953, 960.
[98] John McDowell, 'Virtue and Reason' in *Mind, Value, and Reality* (Harvard University Press 1998) 50.
[99] For example: Michael Oakeshott, 'Rationalism in Politics' (1947) 1 Cambridge Journal 81.

Instead it is the organic growth of institutions characterised by 'attentive silence',[100] more concerned with understanding and responding to particular abuses than with grandiose initiatives from the institutional centre. In an era of international human rights law, this different ideal can seem perversely out of touch with the times.

6. Humility and the UNCRPD

6.1 The compatibility of the UNCRPD with humility

The previous section argues for legislative humility, awareness of the limits of the law; and notes an apparent conflict between such humility and international human rights instruments. In the case of the UNCRPD, this conflict could appear stark. After all, the Convention attempts to 'to promote, protect and ensure the full and equal enjoyment of all human rights and fundamental freedoms by all persons with disabilities'[101] across the world. This could be read as using the law to impose a new vision of society on the grandest possible scale. Furthermore, the scholarly reception of the Convention sometimes seems to reinforce this interpretation. It has almost universally hailed the Convention as a 'paradigm shift' in the relationship between disabled people and the law. This language captures something of the hope that attends the UNCRPD; but, where humility is concerned, it can be dangerous. It can gloss over important limits to understanding. Nowhere in the world is, on even the most generous interpretation, close to the sort of equality that the Convention promises. Accordingly, there is not yet a single example of how to achieve this equality or any concrete details of how it might work in any given society. In these circumstances, rhetoric about implementing a new paradigm too easily implies that the Convention has the answers. As with the Select Committee report into the MCA, the picture is of making a wayward culture align with a clearly established rule; and, once again, that is too optimistic. A more humble appraisal would be that, instead of having answers, the Convention poses the world a set of vitally important questions. This humble appraisal should not be mistaken for despair. Just because the questions posed by the Convention are not yet answered does not mean that they will not be in the future. Good answers to any question, however, are only usually found by those who actively search for them, and only

[100] Simone Weil, 'Human Personality' in Siân Miles (ed), *Simone Weil: An Anthology* (first published 1943, R Rees tr, Penguin 2005) 69, 73.
[101] Art 1.

those who admit that they do not yet know the details of how to make the promise of the Convention real are likely to join this search.

Worries about the language of 'paradigm shifts' are not enough to show that the UNCRPD is necessarily incompatible with legislative humility. In fact, the conflict between the two is more apparent than real, for the UN does not have the institutional machinery of a state. The Convention establishes a Committee on the Rights of Persons with Disabilities that can 'consider' allegations of breaches by states that are party to its Optional Protocol, but even this power is relatively weak. International human rights bodies depend on domestic systems to enforce their decisions, but states do not generally perceive these decisions to be binding, and most are not implemented.[102] The Convention operates in a different context to domestic law, and this is significant. There is no danger of it being used to attempt to directly impose a new vision of society. It is primarily expressive, not directive:[103] a collection of 'standards and values' that can contribute to legal and social change in state parties.[104] Responsiveness is built into its institutional context. This is an example of a point made in section two, humility is especially needed by the powerful. UN human rights bodies do not have the power of states, so humility makes different, lesser, demands on them.[105]

6.2 Fostering humility with the UNCRPD

Beyond mere compatibility, the UNCRPD could help to foster humility when responding to suspected abuse. This might be a surprising claim. After all, the Convention does not mention 'humility', but 'autonomy' is one of its general principles,[106] and Article 16(1) requires 'measures to protect persons with disabilities'. The treaty seems to reinforce the

[102] Rosanne van Alebeek and André Nollkaemper, 'The Legal Status of Decisions by Human Rights Treaty Bodies in National Law' in Helen Keller and Geir Ulfstein (eds), *UN Human Rights Treaty Bodies: Law and Legitimacy* (Cambridge University Press 2012) 356.

[103] Oliver Lewis, 'The Expressive, Educational and Proactive Roles of Human Rights: An Analysis of the United Nations Convention on the Rights of Persons with Disabilities' in Bernadette McSherry and Penelope Weller (eds), *Rethinking Rights-Based Mental Health Laws* (Hart 2010) 97.

[104] Peter Bartlett, 'Implementing a Paradigm Shift: Implementing the Convention on the Rights of Persons with Disabilities in the Context of Mental Disability Law' in *Torture in Healthcare Settings: Reflections on the Special Rapporteur on Torture's 2013 Thematic Report* (Centre for Human Rights and Humanitarian Law 2014) 170.

[105] This does not show that UN human rights bodies *cannot* act in a way that is incompatible with humility, only that the Convention does not *require* it. For some worrying recent developments in human rights monitoring see David McGrogan, 'Human Rights Indicators and the Sovereignty of Technique' (2016) 27 European Journal of International Law 385.

[106] Art 3.

established binary between autonomy and protection, not to challenge it. A careful reading, however, suggests that the Convention could help to foster humility about both knowledge and action. In both cases, this is because of the way that it treats autonomy. It does not treat autonomy as an objective feature of the person that professionals can promote by deploying their 'empowering' expertise. Instead, it makes a link between self-determination and being able to participate, and this undermines any easy assumption that professional expertise can be trusted.

The UNCRPD's potential to promote humility about knowledge becomes apparent when some passages on participation are examined. The preamble affirms that 'persons with disabilities should have the opportunity to be actively involved in decision-making processes about policies and programmes, including those directly concerning them'. Similarly, Article 3 makes 'full and effective participation and inclusion in society' a general principle of the entire Convention, and Article 12(4) requires that any safeguards relating to the exercise of legal capacity respect the 'rights, will and preferences of the person'. All of these Articles emphasise the perspective and expertise of the person with disabilities; and when the person's own perspective is in view, then the relative limits of expert knowledge are more obvious. When limits are more obvious, humility is likely to be easier. It is hard to imagine professionals taking these parts of the Convention seriously but nevertheless removing P in *Somerset* from her family, over her clear objections, without first fully investigating the cause of her bruising. If, then, the UNCRPD influences domestic systems and eventually individual practice, it could nurture epistemic humility. It will, however, only do so to a point. The emphasis on participation draws attention to the limits of what a professional can know, but all human knowledge is limited. There will also be limits to what the disabled person, or the disabled person and an expert working together, knows. Humility about knowledge demands more than the Convention.

The UNCRPD may also foster humility about action. Article 16, which requires appropriate measures be taken to prevent and respond to abuse, parallels the rest of the Convention by emphasising participation over expert intervention. In particular, Article 16(2) requires 'assistance and support for persons with disabilities and their families and caregivers, including through the provision of information and education'. The ideal here is not of a powerful expert intervening to 'save' a victim who has no agency, but of equals working together to avoid abusive situations from developing. By acknowledging that those at risk

of abuse still have agency and power, the Convention, once again, makes the limits to expert power more obvious. Furthermore, this extends, beyond prevention, to responses to abuse. Article 16(4) states:

> States Parties shall take all appropriate measures to promote the physical, cognitive and psychological recovery, rehabilitation and social reintegration of persons with disabilities who become victims of any form of exploitation, violence or abuse, including through the provision of protection services. Such recovery and reintegration shall take place in an environment that fosters the health, welfare, self-respect, dignity and autonomy of the person.

This part of the Convention could have made a difference to TB. To merely remove her from the house that she shared with her abusive husband, without providing further support, falls far short of these requirements. The Convention gives a substantive, albeit general, account of what a victim of abuse might require; and this makes it more difficult to ignore the limits of what a single intervention can achieve. Once again, however, humility demands more than the Convention. The Convention emphasises the limits of expert action; but there will also be limits to what the disabled person, or the disabled person and an expert working together, can achieve in any given circumstances.

The UNCRPD could help to foster humility about both knowledge and action when responding to abuse, but humility demands more than the Convention does. This is not surprising. The UNCRPD, although remarkably idealistic, is the product of political negotiation; but humility is something more rarefied, the product of ethical and theological reflection. There is a need for both of these things. Indeed, in the long term, one of the Convention's strengths may be that the hard questions that it poses prompt reflection even on the limits of human rights language. Sometimes, as when Onazi reads the Convention through Weil's concept of attention,[107] this reflection can be deeply productive. It is also a reflection on limits; and so another, albeit unintentional, way that the Convention can foster humility.

[107] Oche Onazi, '[Disability] Justice Dictated by the Surfeit of Love: Simone Weil in Nigeria' (2016) 28 Law and Critique 1.

7. Conclusion

Being humble is difficult. Medieval monks considered it difficult enough to require detailed guides, and it is no easier today. As Grenberg observes, 'to handle the fact of one's limit in a way that is admirable or exemplary is a slippery, difficult thing to do'.[108] Anyone theorising about humility is well advised to remember one slippery difficulty in particular. The complexity of the human world means we must be content, as Aristotle says, with theories that are, at best, true 'roughly and in outline'.[109] In other words, social theories have limits, so no theorist of humility should expect too much from their work. The concept of humility will not end the abuse of disabled people or dispel all related ethical dilemmas. It will not even clear up all of the associated conceptual tangles. As Clement says of 'Humilitie', George Herbert's allegorical poem, 'there is no decisive victory that would render Humilitie the hero'.[110]

Although humility cannot win a decisive victory against abuse, there is reason to think that it might improve responses to abuse. This chapter shows how it could improve responses at an individual level; humility about knowledge and action can guide a responder when the principles of protection and respect for autonomy alone are inadequate. This has implications for legislation. If non-discrimination requires effective general responses to abuse, and effective responses to abuse require humility, then legislation should foster humility. Furthermore, the law has limits, and the legislator needs humility at least as much as the social worker. Finally, humility highlights the importance of participation as a general principle of the UNCRPD, and this underlines the differences between the Convention and more established approaches. Beyond that, humility can foster awareness of the limits of the treaty itself, and encourage treating it as a set of questions, not answers.

This chapter has many limits. It offers no proof that humility is important when responding to abuse, whether from first principles or by empirical study. Instead, it only offers a set of illustrative examples from one jurisdiction. It does not suggest how humility can be fostered when responding to abuse. Instead, it makes the preliminary point that, if possible, humility should be fostered. Finally, it does not make any detailed legislative proposals. Instead, it only indicates that legislative reform should both pay attention to humility and be humble.

[108] (n 9) 6.
[109] (n 89) 1094b13-28.
[110] (n 11) 25.

Some of these limits could be overcome by further work. Such work, however, would only be valuable if the central argument here, that humility is important when responding to abuse, is plausible. If it is not, then at least the author, as Bernard says, 'will have nothing about which to be proud'.[111]

[111] (n 22) 28.

Chapter Eight

Conclusion

It seems incongruous to follow a chapter arguing for humility with one that claims the author is particularly original. Nevertheless, the genre this work appears within, and the institutions that make it possible, have their own demands; so the first part of this conclusion discusses the originality of the thesis. The second looks forward, to future avenues of research.

1. Originality

Reflecting the 'by publication' format, original contributions are scattered throughout the chapters of this thesis, rather than concentrated in one place. All of the substantive chapters, from two to seven, examine material or develop arguments that are in some way new to the legal literature.

Dual models of autonomy are increasingly influential, but chapter two questions whether they can plausibly distinguish between when it is and is not acceptable to act against a person's expressed wishes while remaining value-neutral. This broad question has been previously addressed in the literature,[1] but the chapter is the first systematic examination of all the possible standards of 'reflection' in a dual model, and methods of assessing it, from the perspective of a person directly trying to apply such a model. This allows it to focus attention more directly on the practical problems with trying to ground value-neutrality in dual models of autonomy.

Chapter three, which is also a paper, examines the relationship between autonomy (in the legal sense) and mental capacity in the courts. It is not the first treatment of this topic.[2] It is,

[1] For example: Jillian Craigie, 'Capacity, Value Neutrality and the Ability to Consider the Future' (2013) 9 International Journal of Law in Context 4; Fabian Freyenhagen and Tom O'Shea, 'Hidden Substance: Mental Disorder as a Challenge to Normatively Neutral Accounts of Autonomy' (2013) 9 International Journal of Law in Context 53; Natalie Banner, 'Can Procedural and Substantive Elements of Decision-making be Reconciled in Assessments of Mental Capacity?' (2013) 9 International Journal of Law in Context 71.

[2] For example: Mary Donnelly, *Healthcare Decision-Making and the Law: Autonomy, Capacity, and the Limits of Liberalism* (Cambridge University Press 2010); John Coggon, 'Mental Capacity Law, Autonomy, and Best Interests: An Argument for Conceptual and Practical Clarity in the Court of Protection' (2016) 24 Medical Law Review 396.

however, original in two regards. It is the first presentation of the three conflicting judicial accounts of the relationship between autonomy and capacity; and it shows that although judge's rhetoric is contradictory, a coherent account of the relevant domestic law can be constructed.

Chapter four draws together the philosophical analysis in chapter two and the legal analysis in chapter three. Because the earlier chapters give relatively full accounts of 'autonomy' in dual models and in recent law it can compare the two more systematically than has previously been done. This leads to the conclusion that dual model autonomy is of less importance than has been assumed in some areas of legal scholarship, although it may be a significant factor in some best interests decisions. Furthermore, this finding allows the development of the novel four-stage justification for actions against a person's expressed wishes: one that requires prospective harm, false belief, necessity, and proportionality.

Chapter five is in two parts. The first examines alternative models of autonomy, under the broad banner of 'relational autonomy'. The link between these models and mental capacity law has previously been addressed by others.[3] The relational autonomy literature, however, often focuses on the promotion of autonomy, not the justification of actions against a person's expressed wishes. By focussing only on the latter issue, the chapter shows that relational models cannot plausibly address the problems raised in chapter two. This conclusion grounds the suggestion that, contrary to the habitual language of both judges and philosophers, there is no 'concept of autonomy' that can determine when it is permissible to act against another person's wishes in their best interests. The chapter suggests an original response to this situation. It draws attention to 'the golden age of female philosophy', writers who are almost entirely neglected in mental disability law, and argues that their methodologies could be useful to the ethics of the field.

Chapter six, which is also a paper, argues that Meno's paradox has implications for mental capacity law. In doing so, it draws attention to areas that have been relatively neglected in the literature. A person's wishes and feelings are an established part of deciding their best

[3] For example: Lucy Series, 'Relationships, Autonomy and Legal Capacity: Mental Capacity and Support Paradigms' (2015) 40 International Journal of Law and Psychiatry 80; Camillia Kong, *Mental Capacity in Relationship* (Cambridge University Press 2017).

interests,[4] but less attention has been paid to their view at the stage of capacity assessment. Although work on advance directives for people with bipolar disorder indicates a similar direction of travel,[5] this paper is the first to argue that there is a general duty to obtain the person's own view on their mental capacity and, in the first instance, treat it as of evidence of what they assert. In addition, the chapter highlights intrinsic limits to advocacy; and shows that even when acting against someone's wishes might be justified, it triggers a duty to support them to avoid such situations in the future.

Finally, chapter seven, another paper, argues that humility should be given more attention than it generally is when the abuse of an adult with mental disabilities is suspected. The account of humility it relies on is not novel, but it has not previously been used to analyse these situations. Doing so gives a better understanding than relying on the customary dichotomy between autonomy and protection alone, for it makes the prospective intervener a direct object of evaluation. Perhaps as importantly, the chapter also serves as a 'proof of concept'. It shows that the ethics of mental disability law can be fruitfully done using the methods of the 'golden age of female philosophy': attending to the genuine moral difficult of these situations; paying close attention to detailed examples; and drawing productive concepts from the history of ideas.

2. Future research

Applying the 'golden age' methodologies to mental disability law also suggests future avenues of research. Most obviously, humility-based analysis might be usefully extended beyond situations of suspected abuse. For instance, it may help to understand the sometimes fraught relationships between professionals and the families of people with mental disabilities, an issue that chapter seven only briefly mentioned. This methodology also has less obvious extensions. For instance, the focus in this thesis has been on potential justifications for not following a person's expressed wishes. This limits its scope to situations where a person's wishes are understood. The question of interpreting a person's wishes has largely been left to one side. Interpretation is, however, a huge topic in its own right, and one that golden age authors have much to contribute to. In particular, Midgley

[4] Mental Capacity Act 2005 s4(6)(a).

[5] Tania Gergel and Gareth S Owen, 'Fluctuating Capacity and Advance Decision-Making in Bipolar Affective Disorder — Self-binding Directives and Self-Determination' (2015) 40 International journal of Law and Psychiatry 92.

presents an account of interpretation that treats the 'scientific', 'historical', and 'phenomenological' categories as equally important and not practically reducible to one another.[6] In everyday life, people use all of these modes. For example, if someone is grumpy with their spouse, then others might explain it by referring to their physical exhaustion, to the terrible example that their parents set, or to the frustration they feel at not getting an expected promotion at work. Something similar seems to happen in mental capacity cases. A person's behaviour is explained sometimes in terms of disorder, sometimes as a reaction to life experiences, and sometimes in terms of what it is like to them. Using Midgley's categories to examine the case law of the Court of Protection may help to understand how interpretation works in practice.

Midgley and her friends deliberately broadened ethical language while paying close attention to human relationships. This is a methodology with vast potential, but one particular concept might benefit mental disability law: loyalty. O'Neill has argued that healthcare workers must be trustworthy,[7] a point that is certainly generalizable to social care. There may, however, be limits to how much trust can ever be achieved in the relationship between a 'professional' and a 'service-user'. Some forms of support might only be possible if the person supported knows that loyalty to them trumps loyalty to a line-manager, to a pay-packet, or even to a professional identity; and this may present formidable barriers to good institutional design. Placing loyalty in the centre of an analysis may help to identify these barriers. Such work need not start from nothing. Loyalty is addressed in detail by Royce, who was discussed in chapter two as one of the sources of modern dual models.[8] Royce's philosophy is not reducible to a dual model alone. It requires individuals to have 'found out some cause, far larger than our individual selves',[9] and directly contrasts this with excessive individualism.[10]

There are also several areas of law that this thesis can provide a platform for further engagement with. Domestically, both the Mental Health Act 1983 and the law pertaining to children allow decisions to be made against a person's wishes for their benefit. These could both be fruitfully put into dialogue with the four-stage justification in chapter four. This

[6] Mary Midgley, *The Ethical Primate: Humans, Freedom and Morality* (Routledge 1994) ch 6.
[7] Onora O'Neill, *Autonomy and Trust in Bioethics* (Cambridge University Press 2002) ch 6.
[8] Chapter two, section 2; Josiah Royce, *The Philosophy of Loyalty* (MacMillan 1911).
[9] *Ibid* 170.
[10] *Ibid* 98.

dialogue could work in both directions. On one hand, these areas of practice can be used to test the justification: can it adequately explain practice? On the other hand the justification can be used to test law and practice: if they do diverge from it, then are those divergences defensible? This process would also allow the account of the law in chapter three to be given additional depth; and may provide a novel perspective on 'fusion' laws, which subsume mental health law into mental capacity law, such as the Mental Capacity Act (Northern Ireland) 2016.

Finally, the UN Convention on the Rights of Persons with Disabilities (UNCRPD) has produced considerable uncertainty in this area of law.[11] Although the Convention is mentioned in almost every chapter of this thesis, it is not dealt with systematically. This work may, however, contribute towards one half of such a systematic evaluation. Understanding what the Convention demands of domestic law requires a full understanding of the law, its justification, the Convention, and its justification. This thesis contributes towards understanding the first two of these: domestic law and its justification. It also, however, repeatedly stresses the difficulty and subtlety of real practice. In particular, it shows that domestic law treats a person's expressed wishes in subtle ways that can appear contradictory without necessarily being so,[12] and that the ethics of justification are difficult and multifaceted in every single case that a person's wishes are not respected.[13] Beyond that, it argues that the relationship between the two, law and ethics, introduces further levels of complexity: law does not simply instantiate ethical rules, but neither is ethics mere excuses for existing practices.[14] It is to be expected that the other half of the picture, the Convention and its justification, is similarly complex. To understand what the Convention demands, what prominent interpretations of it demand,[15] and how they might be justified, all of these things must be put in dialogue with the real situations, as this thesis has attempted to do with domestic law. Only when both sides, domestic law and the

[11] Convention on the Rights of Persons with Disabilities (adopted 13 December 2006, entered into force 3 May 2008) 2515 UNTS 3.
[12] Chapter three.
[13] Chapter four, subsection 3.5.
[14] Chapter seven, subsection 5.3.
[15] For instance, 'General Comment No 1' states that 'the "best interpretation of will and preferences" must replace the "best interests" determinations'. Does the 'best' of 'best interpretation' refer to something like dual model autonomy, what the person *really* wants? It is easy to assume so, but it may mean something like an 'interpretation guided by human rights norms' given that supported decision-making must 'respect human rights norms'. To use the terminology of chapter two, the General Comment explicitly uses 'external standards'. UN Committee on the Rights of Persons with Disabilities, 'General Comment No 1' (2014) UN Doc CRPD/C/GC/1 para 21, 29.

Convention, are given the fullest account of what they demand and the most charitable account of why they demand it can the degree of conflict between the two be properly assessed, and only then can any conflict be truly settled. That, however, is work for another day.

Bibliography

Alebeek RV and Nollkaemper A, 'The Legal Status of Decisions by Human Rights Treaty Bodies in National Law' in Helen Keller and Geir Ulfstein (eds), *UN Human Rights Treaty Bodies: Law and Legitimacy* (Cambridge University Press 2012) 356

Allen N, 'Is Capacity "In Sight"?'(2009) 19 Journal of Mental Health Law 165

Anscombe GEM, 'Does Oxford Moral Philosophy Corrupt Youth?' (1957) 57 The Listener 266

--*Intention* (first published 1957, 2nd ed, Harvard University Press 2000)

--'Modern Moral Philosophy' (1958) 33 Philosophy 1

Aquinas T, *Summa Theologica* (first published 1274, Fathers of the English Dominican Province tr, Catholic Way Publishing 1947)

Arendt H, *The Life of the Mind* (Harcourt 1971)

Aristotle, *Nicomachean Ethics* (first published approximately 322BC, Christopher Rowe tr, Oxford University Press 2002)

Atkins K, 'Autonomy and the Subjective Character of Experience' (2000) 17 Journal of Applied Philosophy 71

-- 'Narrative Identity, Practical Identity and Ethical Subjectivity' (2004) 37 Continental Philosophy Review 341

Augustine, *Nature and Grace* in JE Rotelle (ed), *The Works of Saint Augustine Volume 23* (first published 415, RJ Teske tr, New City Press 1997) 197

Bach M and Kerzner L, *A New Paradigm for Protecting Autonomy and the Right to Legal Capacity* (Law Commission of Ontario 2010)

Banner NF, 'Can Procedural and Substantive Elements of Decision-making be Reconciled in Assessments of Mental Capacity?' (2013) 9 International Journal of Law in Context 71

--and Szmukler G, '"Radical Interpretation" and the Assessment of Decision-Making Capacity' (2013) 30 Journal of Applied Philosophy 379

Bartlett P, 'The United Nations Convention on the Rights of Persons with Disabilities and Mental Health Law' (2012) 75 Modern Law Review 752

--'Implementing a Paradigm Shift: Implementing the Convention on the Rights of Persons with Disabilities in the Context of Mental Disability Law' in *Torture in Healthcare Settings: Reflections on the Special Rapporteur on Torture's 2013 Thematic Report* (Centre for Human Rights and Humanitarian Law 2014) 170

Bauman CW and others, 'Revisiting External Validity: Concerns about Trolley Problems and Other Sacrificial Dilemmas in Moral Psychology' (2014) 8 Social and Personality Psychology Compass 536

Benson P, 'Free Agency and Self-Worth' (1994) 91 The Journal of Philosophy 650

--'Feminist Intuitions and the Normative Substance of Autonomy' in James Stacey Taylor (ed), *Personal Autonomy: New Essays on Personal Autonomy and its Role in Contemporary Moral Philosophy* (Cambridge University Press 2005) 124

Beresford P and others, *From Mental Illness to a Social Model of Madness and Distress* (Shaping our Lives 2016)

Berlin I, 'Two Concept of Liberty' in *The Proper Study of Mankind: An Anthology of Essays* (first published 1958, Pimlico 1998) 191

Bernard of Clairvaux, *The Steps of Humility and Pride* (first published 1120, M Ambrose Conway tr, Cistercian Publications 1973)

Bortolotti L, *Irrationality* (Polity Press 2015)

--and Antrobus M, 'Costs and Benefits of Realism and Optimism' (2015) 28 Current Opinion in Psychiatry 194

Boyd CA, 'Pride and Humility: Tempering the Desire for Excellence' in Kevin Timpe and Craig A Boyd (eds), *Virtues and their Vices* (Oxford University Press 2014) 245

Bratman ME, *Intention, Plans, and Practical Reason* (Harvard University Press 1987)

-- 'Identification, Decision, and Treating as a Reason' in *Faces of Intention: Selected Essays on Intention and Agency* (Cambridge University Press 1999) 185

--*Structures of Agency: Essays* (Oxford University Press 2007)

Burch M, 'Autonomy, Respect, and the Rights of Persons with Disabilities in Crisis' [2016] 34 Journal of Applied Philosophy 389

Button M, '"A Monkish Kind of Virtue?" For and Against Humility' (2005) 33 Political Theory 840

Calcott B, 'Wimsatt and the Robustness Family' (2011) 26 Biology and Philosophy 284

Carlsson R and Agerström J, 'A Closer Look at the Discrimination Outcomes in the IAT Literature' (2016) 57 Scandinavian Journal of Psychology 278

Case P, 'Dangerous Liaisons? Psychiatry and Law in the Court of Protection — Expert Discourses of 'Insight' (and 'Compliance')' (2016) 24 Medical Law Review 360

--'Negotiating the Domain of Mental Capacity: Clinical Judgement or Judicial Diagnosis?' (2016) 16 Medical Law International 174

Cave E, 'Determining Capacity to Make Medical Treatment Decisions: Problems Implementing the Mental Capacity Act 2005' (2014) 36 Statute Law Review 86

--and Tan J, 'Severe and Enduring Anorexia Nervosa in the Court of Protection in England And Wales' (2017) 23 International Journal of Mental Health and Capacity Law 4

Chesterton GK, *Orthodoxy* in David Dooley (ed), *GK Chesterton: Collected Works Volume I* (first published 1908, Ignatius Press 1986) 209

Christman J, *The Politics of Persons: Individual Autonomy and Socio-Historical Selves* (Cambridge University Press 2009)

Clement J, *Reading Humility in Early Modern England* (Routledge 2015)

Clough B, 'People Like That: Realising the Social Model in Mental Capacity Jurisprudence' (2014) 23 Medical Law Review 53

--'What About Us? A Case for Legal Recognition of Interdependence in Informal Care Relationships' (2014) 36 Journal of Social Welfare and Family Law 129

--'Anorexia, Capacity, and Best Interests: Developments in the Court Of Protection since the Mental Capacity Act 2005' (2016) 24 Medical Law Review 434

Coggon J, 'Varied and Principled Understandings of Autonomy in English Law: Justifiable Inconsistency or Blinkered Moralism?' (2007) 15 Health Care Analysis 235

-- 'Mental Capacity Law, Autonomy, and Best Interests: An Argument for Conceptual and Practical Clarity in the Court of Protection' (2016) 24 Medical Law Review 396

--and Miola J, 'Autonomy, Liberty, and Medical Decision-Making' (2011) 70 Cambridge Law Journal 523

Cooper JE, *Secular Powers: Humility in Modern Political Thought* (University of Chicago Press 2013)

Coulehan J, 'A Gentle and Humane Temper: Humility in Medicine' (2011) 54 Perspectives in Biology and Medicine 206

Craigie J, 'Capacity, Value Neutrality and the Ability to Consider the Future' (2013) 9 International Journal of Law in Context 4

DasGupta S, 'Narrative Humility' (2008) 371 The Lancet 980

Davidson D, 'Psychology as Philosophy', in *Essays on Actions and Events* (Oxford University Press 2001) 229

Davy L, 'Philosophically Inclusive Design: Intellectual Disability and the Limits of Individual Autonomy in Moral and Political Theory' (2015) 20 Hypatia 132

Dawson J, 'A Realistic Approach to Assessing Mental Health Laws' Compliance with the UNCRPD' (2015) 40 International Journal of Law and Psychiatry 70

Department for Constitutional Affairs, *The Mental Capacity Act 2005 Code of Practice* (The Stationery Office 2007)

Department of Health, *Care and Support Statutory Guidance* (Department of Health 2014)

--*Mental Health Act 1983: Code of Practice* (The Stationery Office 2015)

--*The Seventh Year of the Independent Mental Capacity Advocacy (IMCA) Service* (Department of Health 2015)

Descartes R, *The Passions of the Soul* in *The Philosophical Writings of Descartes Volume 1* (first published 1649, John Cottingham and others tr, Cambridge University Press 1985) 325

Dhanda A, 'Legal Capacity in the Disability Rights Convention: Stranglehold of the Past or Lodestar for the Future?' (2007) 34 Syracuse Journal of International Law and Commerce 429

Diesfeld K and Sjöström S, 'Interpretive Flexibility: Why Doesn't Insight Incite Controversy in Mental Health Law?' (2007) 25 Behavioural Sciences and the Law 85

Dodds S, 'Choice and Control in Feminist Bioethics' in Catriona Mackenzie and Natalie Stoljar (eds), *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self* (Oxford University Press 2000) 213

Donnelly M, 'Best Interests, Patient Participation and the Mental Capacity Act 2005' (2009) 17 Medical Law Review 1

--'Capacity Assessment under the Mental Capacity Act 2005: Delivering on the Functional Approach?' (2009) 29 Legal Studies 464

--Healthcare Decision-Making and the Law: Autonomy, Capacity, and the Limits of Liberalism (Cambridge University Press 2010)

Dunn MC and others, 'To Empower or to Protect? Constructing the 'Vulnerable Adult' in English Law and Public Policy' (2008) 28 Legal Studies 234

Dunning D, *Self-Insight: Roadblocks and Detours on the Path to Knowing Thyself* (Taylor and Francis 2005)

Dworkin G, *The Theory and Practice of Autonomy* (Cambridge University Press 1988)

Dworkin R, 'Liberalism' in *A Matter of Principle* (Harvard University Press 1985) 191

Ehrlinger J and others, 'Understanding Overconfidence: Theories of Intelligence, Preferential Attention, and Distorted Self-Assessment' (2016) 63 Journal of Experimental Social Psychology 94

Ellis JW, 'Decisions by and for People with Mental Retardation: Balancing Considerations of Autonomy and Protection' (1992) 37 Villanova Law Review 1779

Emmett C and others, 'Homeward Bound or Bound for a Home? Assessing the Capacity of Dementia Patients to Make Decisions about Hospital Discharge: Comparing Practice with Legal Standards' (2013) 36 International Journal of Law and Psychiatry 73

Essex Autonomy Project, 'Achieving CRPD Compliance: Is the Mental Capacity Act of England and Wales Compatible with the UN Convention on the Rights of Persons with Disabilities? If Not, What Next?' (University of Essex 2014)

Fanning J, 'Continuities of Risk in the Era of the Mental Capacity Act' (2016) 24 Medical Law Review 415

Fine G, 'Inquiry in the Meno' in Richard Kraut (ed), *The Cambridge Companion to Plato* (Cambridge University Press 1992)

Flynn E and Arstein-Kerslake A, 'The Support Model of Legal Capacity: Fact, Fiction, or Fantasy?' (2014) 32 Berkeley Journal of International Law 124

--'Legislating Personhood: Realising the Right to Support in Exercising Legal Capacity' (2014) 10 International Journal of Law in Context 81

Foot P, 'The Problem of Abortion and the Doctrine of the Double Effect' (1967) 5 Oxford Review 1

--*Natural Goodness* (Oxford University Press 2001)

Frankfurt HG, 'Freedom of the Will and the Concept of a Person' (1971) 68 the Journal of Philosophy 5

--*On Bullshit* (first published 1986, Princeton University Press 2005)

Freyenhagen F and O'Shea T, 'Hidden Substance: Mental Disorder as a Challenge to Normatively Neutral Accounts of Autonomy' (2013) 9 International Journal of Law in Context 53

Friedman MA, 'Autonomy and the Split-Level Self' (1986) 24 Southern Journal of Philosophy 19

--*Autonomy, Gender, Politics* (Oxford University Press 2003)

--'Review: Christman, John. The Politics of Persons: Individual Autonomy and Socio-Historical Selves' (2011) 121 Ethics 424

Fulford KWM and others, 'Concepts of Disease and the Abuse of Psychiatry in the USSR' (1993) 162 British Journal of Psychiatry 801

--*Oxford Textbook of Philosophy and Psychiatry* (Oxford University Press 2006)

Garnett M, 'Taking the Self out of Self-Rule' (2013) 16 Ethical Theory and Moral Practice 21

--'The Autonomous Life: A Pure Social View' (2014) 92 Australasian Journal of Philosophy 143

--'Agency and Inner Freedom' (2017) 51 Noûs 3

Gergel T and Owen GS, 'Fluctuating Capacity and Advance Decision-Making in Bipolar Affective Disorder — Self-binding Directives and Self-Determination' (2015) 40 International journal of Law and Psychiatry 92

Glover J, *Alien Landscapes? Interpreting Disordered Minds* (Harvard University Press 2014)

Glover-Thomas N, 'The Age of Risk: Risk Perception and Determination Following the Mental Health Act 2007' (2011) 19 Medical Law Review 581

Gooding P, 'Navigating the 'Flashing Amber Lights' of the Right to Legal Capacity in the UNCRPD: Responding to Major Concerns' (2015) 15 Human Rights Law Review 45

--and Flynn E, 'Querying the Call to Introduce Mental Capacity Testing to Mental Health Law: Does the Doctrine of Necessity Provide an Alternative?' (2015) 4 Laws 245

Green B and others, *Independent Review of Deaths of People with a Learning Disability or Mental Health Problem in Contact with Southern Health NHS Foundation Trust April 2011 to March 2015* (Mazars 2015)

Grenberg J, *Kant and the Ethics of Humility* (Cambridge University Press 2005)

Hämäläinen N, 'What is a Wittgensteinian Neo-Platonist? – Iris Murdoch, Metaphysics and Metaphor (2014) 43 Philosophical Papers 191

Hansen C, 'Wuwei: Taking No Action' in Antonio S Cua (ed), *Encyclopaedia of Chinese Philosophy* (Routledge 2003) 786

Harrington J, *Towards a Rhetoric of Medical Law* (Taylor and Francis 2017)

Hedley M, *The Modern Judge: Power, Responsibility and Society's Expectations* (LexisNexis 2016)

Herring J, 'Losing it? Losing what? The Law and Dementia' (2009) 21 Child and Family Law Quarterly 3

--and Wall J, 'Autonomy, Capacity and Vulnerable Adults: Filling the Gaps in the Mental Capacity Act (2015) 35 Legal Studies 698

Heslop P and others, *The Confidential Inquiry into Premature Deaths of People with Learning Disabilities (CIPOLD)* (Norah Fry Research Centre 2013)

Hewson B, '"Neither Midwives nor Rainmakers" - Why DL Is Wrong' (2013) 3 Public Law 451

Hope T and others, 'Anorexia Nervosa and the Language of Authenticity' (2011) 41 Hastings Centre Report 19

Huemer M, *The Problem of Political Authority: An Examination of the Right to Coerce and the Duty to Obey* (Palgrave Macmillan 2013)

Hughes K and others, 'Prevalence and Risk of Violence against Adults with Disabilities: A Systematic Review and Meta-Analysis of Observational Studies' (2012) 379 The Lancet 1621

Humboldt WV, *The Limits of State Action* (first published 1850, JW Burrow tr, Cambridge University Press 1969)

Hume D, *Dialogues Concerning Natural Religion* (first published 1779, Dorothy Coleman ed, Cambridge University Press 2007)

Hursthouse R, *On Virtue Ethics* (Oxford University Press 2001)

James W, *The Principles of Psychology* (Henry Holt and Company 1890)

Johnston C and others, 'Patient Narrative: An "On-Switch" for Evaluating Best Interests' (2016) 38 Journal of Social Welfare and Family Law 249

Joy-Gaba JA and Nosek BA, 'The Surprisingly Limited Malleability of Implicit Racial Evaluations' (2010) 41 Social Psychology 137

Kain P, 'Kant's Defense of Human Moral Status' (2009) 47 Journal of the History of Philosophy 59

Kant I, *Groundwork of the Metaphysics of Morals* (first published 1785, Mary Gregor tr, Cambridge University Press 1997)

--*Religion Within the Boundaries of Mere Reason* in *Religion and Rational Theology* (first published 1793, Allen W Wood and George Di Giovanni tr, Cambridge University Press 1996)

--*The Metaphysics of Morals* (first published 1797, Mary Gregor tr, Cambridge University Press 1996)

Keywood K, '"I'd Rather Keep Him Chaste." Retelling the Story of Sterilisation, Learning Disability and (Non)Sexed Embodiment' (2001) 9 Feminist Legal Studies 185

Kierkegaard, *Concluding Unscientific Postscript to the Philosophical Crumbs* (first published 1846, Alastair Hannay tr, Cambridge University Press 2009)

Killmister S, 'Autonomy and False Beliefs' (2012) 164 Philosophical Studies 513

--'Autonomy, Liberalism, and Anti-Perfectionism' (2013) 19 Res Publica 353

--'The Woody Allen Puzzle: How "Authentic Alienation" Complicates Autonomy' (2015) 49 Noûs 729

Klancer CH, 'How Opposites (Should) Attract: Humility as a Virtue for the Strong' (2012) 22 The Heythrop Journal 662

Kong C, 'The Convention for the Rights of Persons with Disabilities and Article 12: Prospective Feminist Lessons against the "Will and Preferences" Paradigm' (2015) 4 Laws 709

--*Mental Capacity in Relationship* (Cambridge University Press 2017)

Korsgaard CM, *The Sources of Normativity* (Onora O'Neill ed, Cambridge University Press 1996)

Lane KA and others, 'Implicit Social Cognition and Law' (2007) 3 Annual Review of Law and Social Science 427

Law Commission, *Mental Incapacity* (Law Com No 231, 1995)

--*Mental Capacity and Deprivation of Liberty: A Consultation Paper* (Consultation Paper No 222, 2015)

--*Mental Capacity and Deprivation of Liberty* (Law Com No 372, 2017)

Levin M, 'Conditional Rights' (1989) 55 Philosophical Studies 211

Lewis D, *Counterfactuals* (first published 1973, 2nd ed, Wiley-Blackwell 2001)

--*On the Plurality of Worlds* (Blackwell 1986)

Lewis O, 'The Expressive, Educational and Proactive Roles of Human Rights: An Analysis of the United Nations Convention on the Rights of Persons with Disabilities' in Bernadette McSherry and Penelope Weller (eds), *Rethinking Rights-Based Mental Health Laws* (Hart 2010) 97

Lippitt J, 'Getting the Story Straight: Kierkegaard, MacIntyre and Some Problems with Narrative' (2007) 50 Inquiry 34

MacIntyre A, 'The Intelligibility of Action' in J Margolis (ed), *Rationality, Relativism and the Human Sciences* (Martinus Nijhoff 1986)

--*Three Rival Versions of Moral Enquiry* (Duckworth 1990)

Mackenzie C and Rogers W, 'Autonomy, Vulnerability and Capacity: A Philosophical Appraisal of the Mental Capacity Act' (2013) 9 International Journal of Law in Context 37

Maclean A, *Autonomy, Informed Consent and Medical Law: A Relational Challenge* (Cambridge University Press: Cambridge, 2009)

Martin W and others, 'Achieving CRPD Compliance' (Essex Autonomy Project 2014)

--*Three Jurisdictions Report: Towards Compliance with CRPD Art 12 in Capacity/Incapacity Legislation across the UK* (University of Essex 2016)

Martinuk MJM, 'A Fundamental Orientation to the Good: Iris Murdoch's Influence on Charles Taylor in Mark Luprecht (ed), *Iris Murdoch Connected: Critical Essays on Her Fiction and Philosophy* (University of Tennessee Press 2014) 183

McDowell J, 'Two Sorts of Naturalism' in *Mind, Value, and Reality* (Harvard University Press 1998) 189

--'Virtue and Reason' in *Mind, Value, and Reality* (Harvard University Press 1998) 50

McGrogan D, 'Human Rights Indicators and the Sovereignty of Technique' (2016) 27 European Journal of International Law 385

Melnyk, 'Is there a Formal Argument against Positive Rights?' (1989) 55 Philosophical Studies 205

Midgley M, 'Is "Moral" a Dirty Word?' in *Heart and Mind* (first published 1972, Routledge 2003) 119

--'The Game Game' in *Heart and Mind* (first published 1974, Routledge 2003) 154

-- *Can't We Make Moral Judgements?* (first published 1991, Bloomsbury 2017)

--*The Ethical Primate: Humans, Freedom and Morality* (Routledge 1994)

--'The Golden Age of Female Philosophy' *The Guardian* (London, 28 November 2013) <https://www.theguardian.com/world/2013/nov/28/golden-age-female-philosophy-mary-midgley> accessed 20 November 2017

--'A Golden Manifesto' (2016) 116 Philosophy Now 34

--'A Golden Manifesto, Part II' (2016) 117 Philosophy Now 20

Miles J, 'Family Abuse, Privacy and State Intervention' (2011) 70 Cambridge Law Journal 31

Mill JS, *On Liberty* (first published 1859, David Bromwich and George Kateb eds, Yale University Press 2003)

Moll FL and others, 'Management of Abdominal Aortic Aneurysms: Clinical Practice Guidelines of the European Society for Vascular Surgery' (2011) 41 European Journal of Vascular and Endovascular Surgery S1

Mulhall S and Swift A, *Liberals and Communitarians* (first published 1992, 2nd ed, Blackwell 1996)

Murdoch I, 'Against Dryness' (1961) 16 Encounter 16

--*The Sovereignty of Good* (first published 1970, Routledge 2014)

Nagel T, *The View from Nowhere* (Oxford University Press 1986)

Norden BWV, *Taking Back Philosophy: A Multicultural Manifesto* (Columbia University Press 2017)

Nosek BA and others, 'Implicit Social Cognition: From Measures to Mechanisms' (2011) 15 Trends in Cognitive Sciences 152

Nussbaum MC, 'Aristotle, Nature, and Ethics' in JEJ Altham and Ross Harrison (eds) *World, Mind, and Ethics: Essays on the Ethical Philosophy of Bernard Williams* (Cambridge University Press 1995) 86

Nys T, 'Review: The Politics of Persons: Individual Autonomy and Socio-Historical Selves, by John Christman' (2011) 19 European Journal of Philosophy 474

Oakeshott M, 'Rationalism in Politics' (1947) 1 Cambridge Journal 81

Onazi O, '[Disability] Justice Dictated by the Surfeit of Love: Simone Weil in Nigeria' (2016) 28 Law and Critique 1

O'Neill O, *Autonomy and Trust in Bioethics* (Cambridge University Press 2002)

-- 'The Inaugural Address: Autonomy: The Emperor's New Clothes' (2003) 77 Proceedings of the Aristotelian Society, Supplementary Volumes 1

Oshana M, 'How Much Should We Value Autonomy?' (2003) 20 Social Philosophy and Policy 99

Owen GS and others, 'Mental Capacity, Diagnosis and Insight in Psychiatric In-Patients: A Cross-Sectional Study' (2008) 39 Psychological Medicine 1389

--'Mental Capacity and Decisional Autonomy: An Interdisciplinary Challenge (2009) 52 Inquiry 79

Parens E, 'The Pluralist Constellation' (1995) 4 Cambridge Quarterly of Health Care Ethics 197

Peay J, 'Mental Incapacity and Criminal Liability: Redrawing the Fault Lines?' (2015) 40 International Journal of Law and Psychiatry 25

Penrose B, 'Humility and Understanding' (2010) 39 Philosophical Papers 427

Plato, *Meno* in John M Cooper (ed), *Plato: Complete Works* (first published approximately 385BC, GMA Grube tr, Hackett 1997) 870

--*Republic* in John M Cooper (ed), *Plato: Complete Works* (first published approximately 380BC, GMA Grube and CDC Reeve tr, Hackett 1997) 971

Prall SE, 'The Development of Equity in Tudor England' (1964) 8 The American Journal of Legal History 1

Pritchard-Jones L, 'The Good, the Bad, and the "Vulnerable Older Adult"' (2016) 58 Journal of Social Welfare and Family Law 51

--'"This Man with Dementia"—"Othering" the Person with Dementia in the Court of Protection' (2016) 24 Medical Law Review 518

--'Ageism and Autonomy in Health Care: Explorations through a Relational Lens' (2017) 25 Health Care Analysis 72

Putnam RA, 'Weaving Seamless Webs' in David MacArthur (ed), *Pragmatism as a Way of Life: The Lasting Legacy of William James and John Dewey* (Belknap 2017)

Rawls J, *A Theory of Justice: Revised Edition* (first published 1971, Harvard University Press 1999)

Raz J, *The Morality of Freedom* (Clarendon Press 1986)

Richardson G, 'Balancing Autonomy and Risk: A Failure of Nerve in England and Wales?' (2007) 30 International Journal of Law and Psychiatry 71

Royce J, *The Philosophy of Loyalty* (MacMillan 1911)

Scott D, *Plato's Meno* (Cambridge University Press 2006)

Searle J, 'A Taxonomy of Illocutionary Acts,' in K Gunderson (ed), *Language, Mind and Knowledge* (University of Minnesota Press 1975)

--'What Is an Intentional State?' (1979) LXXXVIII Mind 74

Series L, 'Relationships, Autonomy and Legal Capacity: Mental Capacity and Support Paradigms' (2015) 40 International Journal of Law and Psychiatry 80

--'The Place of Wishes and Feelings in Best Interests Decisions' (2016) 79 Modern Law Review 1101

Sher G, *Beyond Neutrality: Perfectionism and Politics* (Cambridge University Press 1997)

Sjöstrand M and others, 'Paternalism in the Name of Autonomy' (2013) 38 Journal of Medicine and Philosophy 710

Skowron P, 'Evidence and Causation in Mental Capacity Assessments' (2014) 22 Medical Law Review 631

--'The Implications of Meno's Paradox for the Mental Capacity Act 2005' (2016) 24 Medical Law Review 279

--'Humility when Responding to the Abuse of Adults with Mental Disabilities' (2017) 53 International Journal of Law and Psychiatry 102

Stefan S, 'Silencing the Different Voice: Competence, Feminist Theory and Law' (1993) 47 University of Miami Law Review 763

Stoljar N, 'Autonomy and the Feminist Intuition' in Catriona Mackenzie and Natalie Stoljar (eds), *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self* (Oxford University Press 2000) 94

Sunstein CR, 'Problems with Rules' (1995) 83 California Law Review 953

Szmukler G, 'The UN Convention on the Rights of Persons with Disabilities: 'Rights, Will and Preferences' in Relation to Mental Health Disabilities' (2017) International Journal of Law and Psychiatry doi.org/10.1016/j.ijlp.2017.06.003

--and Bach M, 'Mental Health Disabilities and Human Rights Protections' (2015) 2 Global Mental Health e20

Tan JOA and others, 'Competence to Make Treatment Decisions in Anorexia Nervosa: Thinking Processes and Values' (2006) 13 Philosophy, Psychiatry and Psychology 267

Taylor C, *Sources of the Self: The Making of the Modern Identity* (Harvard University Press 1989)

--*The Ethics of Authenticity* (Harvard University Press 1991)

--*A Secular Age* (Harvard University Press 2007)

--'Iris Murdoch and Moral Philosophy' in *Dilemmas and Connections: Selected Essays* (Harvard University Press 2011) 3

Thomas EW, *The Judicial Process: Realism, Pragmatism, Practical Reasoning and Principles* (Cambridge University Press 2005)

Watson G, 'Free agency' in Gary Watson (ed) *Free Will* (Oxford University Press 1982)

Weil S, 'Human Personality' in Siân Miles (ed), *Simone Weil: An Anthology* (first published 1943, R Rees tr, Penguin 2005) 69

Westlund AC, 'Selflessness and Responsibility for Self: Is Deference Compatible with Autonomy?' (2003) 112 The Philosophical Review 483

--'Rethinking Relational Autonomy' (2009) 24 Hypatia 26

White NP, *Plato on Knowledge and Reality* (Hackett 1976)

Wilks I, 'The Debate over Risk-Related Standards of Competence' (1997) 11 Bioethics 5

Williams B, *Ethics and the Limits of Philosophy* (first published 1985, Routledge 2011)

--'Ethical Consistency' in Geoffrey Sayre-McCord (ed), *Essays on Moral Realism* (Cornell University Press 1988) 41

Williams V and others, *Making Best Interests Decisions: People and Processes* (Mental Health Foundation 2012)

Wilson MC and Scior K, 'Attitudes towards Individuals with Disabilities as Measured by the Implicit Association Test: A Literature Review' (2014) 35 Research in Developmental Disabilities 294

Wimsatt WC, *Re-engineering Philosophy for Limited Beings* (Harvard University Press 2007)

Winick BJ, 'The Side Effects of Incompetency Labelling and the Implications for Mental Health Law' (1995) 1 Psychology, Public Policy, and Law 6

Wiseman R, 'Anscombe, Foot, Midgley and Murdoch: A Female Philosophical School?' (Annual Women in the History of Philosophy Lecture, The University of Sheffield, 12 May 2017) < http://www.womeninparenthesis.co.uk/anscombe-foot-midgley-and-murdoch-a-female-philosophical-school/> accessed 20 November 2017

Wittgenstein L, *Philosophical Investigations* (first published 1953, PMS Hacker and Joachim Schlute eds, GEM Anscombe and others tr, 4th ed, Wiley-Blackwell 2009)

--'Philosophy of Psychology – A Fragment' in PMS Hacker and Joachim Schlute (eds), *Philosophical Investigations* (first published 1953, GEM Anscombe and others tr, 4th ed, Wiley-Blackwell 2009)

Wolff J, 'How Can We End the Male Domination of Philosophy?' *The Guardian* (London, 26 November 2013) <https://www.theguardian.com/education/2013/nov/26/modern-philosophy-sexism-needs-more-women> accessed 20 November 2017

Wrigley A, 'The Problem of Counterfactuals in Substituted Judgement Decision-Making' (2011) 28 Journal of Applied Philosophy 169

Xiao Y, 'Virtue Ethics as Political Philosophy: The Structure of Ethical Theory in Early Chinese Philosophy' in Lorraine Besser-Jones and Michael Slote (eds), *The Routledge Companion to Virtue Ethics* (Routledge 2015) 471

Appendix

Published Papers

1. 'The Implications of Meno's Paradox for the Mental Capacity Act 2005' (2016) 24 Medical Law Review 279

2. 'Humility when Responding to the Abuse of Adults with Mental Disabilities' (2017) 53 International Journal of Law and Psychiatry 102

Medical Law Review, Vol. 24, No. 3, pp. 379–395
doi:10.1093/medlaw/fww026

THE IMPLICATIONS OF MENO'S PARADOX FOR THE MENTAL CAPACITY ACT 2005

PAUL SKOWRON*

Centre for Social Ethics and Policy, University of Manchester, Manchester, UK

ABSTRACT

Meno's paradox—which asks 'how will you know it is the thing you didn't know?'—appears in Plato's dialogue of the same name. This article suggests that a similar question arises in some supportive relationships. Attention to this question clarifies one condition necessary to justify making a best interests decisions against someone's will: the decided-for person must be unable to recognise that they have failed to recognise a need. From this condition, two duties are derived: a duty to ensure that someone cannot recognise that they have failed to recognise a need before making a decision against their will; and a duty to provide consensual support to those who have had decisions made against their will, in order to help them to avoid such second-order failures of recognition in the future. The article assesses the Mental Capacity Act 2005 against each of these duties. For each duty, it finds that the Act allows compliance, but does not robustly require it.

KEYWORDS: Advocacy, Care Act, Epistemology, Mental Capacity, Paradox, Decision-Making

I. INTRODUCTION

When, if ever, is it right to make a decision for another person that they object to? The question seems to be intractable. The Mental Capacity Act 2005 (the MCA) provides one set of answers, but the UN Convention on the Rights of

* School of Law, Williamson Building 4.42, University of Manchester, Oxford Road, Manchester, M13 9PL, UK, paul.skowron@postgrad.manchester.ac.uk. I would like to thank to Nicola Glover-Thomas, Kirsty Keywood, Iain Brassington, Laura Pritchard-Jones, and the two anonymous reviewers for their constructive comments on this article. This work was first presented at 'The Mental Capacity Act 2005: Ten Years On' Conference at Liverpool on 9–10 September 2015, which was supported by the Wellcome Trust (Grant No 107461/Z/15/Z). No conflicts of interest to declare.

Persons with Disabilities (the UNCRPD)[1] offers another, apparently incompatible, set.[2] The MCA says that such a decision can be made if it is in someone's best interests and they cannot understand, retain, or use and weigh the relevant information due to 'an impairment of, or a disturbance in the functioning of, the mind or brain'.[3] The UNCRPD, in contrast, has been interpreted to prohibit all 'substituted decision-making': defined as any process in which a third party is appointed against someone's will to make decisions for them; and the decision is made in the person's best interests, rather than according to the best interpretation of their will and preferences.[4] Compelling points have been made against both of these positions, and each has been given increasingly subtle interpretations to accommodate these objections.[5] The resulting dialogue, however, does little to dispel the impression that the core question is intractable; for, on this, neither side has appreciably changed its answer.

This article does not solve the question of whether it is better to make decisions for others in their bests interests or according to the best interpretation of their will and preferences. Instead, it draws attention to a paradox that can occur in supportive relationships. From this paradox, this article shows that best interests decision-making against a person's will, *if* it is justified in a liberal society, must be premised on the decided-for person being unable to recognise that they have failed to recognise a need. From this premise, two duties can be derived; and this allows any particular system that uses a best interests standard, in this case the MCA, to be evaluated against standards that it can be assumed to be implicitly committed to. This is not the same task as assessing the Act against an external normative framework, such as the one that the UNCRPD presents. Nevertheless, it is striking that if these duties were taken seriously, then it would narrow, without entirely closing, the controversial gap between the two systems.

The next section develops an account of the paradox. Section III derives the first duty, the duty of identification, and assesses the MCA against it. Section IV then evaluates the extent to which legal requirements to provide advocacy might require compliance with this duty. Finally, Section V derives the second duty, the duty to support, and assesses the Act against this. For each duty, it is found that the MCA allows compliance,

1 Convention on the Rights of Persons with Disabilities (adopted 13 December 2006, entered into force 3 May 2008) 2515 UNTS 3.

2 Peter Bartlett, 'The United Nations Convention on the Rights of Persons with Disabilities and Mental Health Law' (2012) 75 Modern Law Review 752.

3 MCA 2005 s 1–4. An inability to communicate, under s 3(1)(d), is not relevant to a decision that someone objects to.

4 UN Committee on the Rights of Persons with Disabilities, 'General Comment No 1' (2014) UN Doc CRPD/C/GC/1 para 27.

5 Elionóir Flynn and Anna Arstein-Kerslake, 'The Support Model of Legal Capacity: Fact, Fiction, or Fantasy?' (2014) 32 Berkeley Journal of International Law 124; Wayne Martin and others, 'Achieving CRPD Compliance' (Essex Autonomy Project 2014) 10–13; John Dawson, 'A Realistic Approach to Assessing Mental Health Laws' Compliance with the UNCRPD' (2015) 40 International Journal of Law and Psychiatry 70; Lucy Series, 'Relationships, Autonomy and Legal Capacity: Mental Capacity and Support Paradigms' (2015) 40 International Journal of Law and Psychiatry 80; Piers Gooding, 'Navigating the "Flashing Amber Lights" of the Right to Legal Capacity in the UNCRPD: Responding to Major Concerns' (2015) 15 Human Rights Law Review 45.

but stops some way short of requiring it. In the conclusion, these limited criticisms of the Act are reintegrated into the context of the wider debate.

II. MENO'S PARADOX IN SUPPORTIVE RELATIONSHIPS

Meno's paradox is presented by Plato in the dialogue of the same name. It is stated in two ways: first by Meno and then by Socrates.[6] Socrates' statement of the problem is slightly clearer. He says:

> [A] man cannot search either for what he knows or what he does not know. . . He cannot search for what he knows – since he knows it, there is no need to search – nor for what he does not know, for he does not know what to look for.[7]

A common response to this paradox is to deny that knowledge is as binary as it suggests. So, for instance, Fine interprets Plato's Socrates as proposing a distinction between 'knowledge' and 'true beliefs'.[8] On this interpretation, we may believe things to be true that are actually true without knowing why they are true.[9] These true beliefs then provide us with a starting point from which we can develop a deeper understanding;[10] and this understanding may, in time, develop to the point where we can give a coherent account of why our beliefs are true. It is this ability to explain 'why' that distinguishes knowledge from belief.[11] There is, however, a form of the paradox that this distinction does not resolve.

Meno originally asks Socrates 'how will you know it is the thing you didn't know?'[12] This suggests a particularly difficult problem, the problem of recognition.[13] If knowledge requires being able to give an account of the reasons why a belief is true, then those 'reasons' will inevitably refer to other beliefs that, to count as knowledge, must in turn rely on even deeper reasons. There is no obvious end to this process, and so no certain way of distinguishing between what you know and what you falsely believe that you know.[14] There is always the possibility that something we take to be knowledge is actually premised on a deep unrecognised error. Sometimes, this is directly relevant to supportive relationships.

A paradox of recognition can occur in supportive relationships. This can be shown by an example. Imagine a man called 'John', who has a large (over 9 cm) abdominal aortic aneurysm that is growing rapidly. John wants to live. The diagnosis of such an aneurysm is reliable, they are very likely to rupture, and rupture carries an extremely

6 Plato, 'Meno' in John M Cooper (ed), *Plato: Complete Works* (GMA Grube tr, Hackett 1997) 80d80e.
7 ibid 80e.
8 Gail Fine, 'Inquiry in the Meno' in Richard Kraut (ed), *The Cambridge Companion to Plato* (CUP 1992) 200, 206.
9 Perhaps because we simply accept what an authority, such as a teacher, says. Alasdair MacIntyre, *Three Rival Versions of Moral Enquiry* (Duckworth 1990) 63.
10 Scott distinguishes between knowing 'parrot fashion' and 'synoptically'. Dominic Scott, *Plato's Meno* (CUP 2006) 102–03, 183–84.
11 Plato (n 6) 98.
12 ibid 80d.
13 Nicholas P White, *Plato on Knowledge and Reality* (Hackett 1976) 42–47.
14 Scott (n 10) 83–84.

high risk of death.[15] Assuming, for now, that there are no other complicating factors, it is reasonable to say that John objectively needs a particular form of support: surgery. Further imagine, however, that John does not recognise his need for this support. He is asymptomatic, and simply disbelieves the diagnosis. In these circumstances, those supporting him might want to help him to recognise that if he wishes to live, he needs surgery. This is when the paradox appears. If John accepts this help—a second-order support to recognise his need for first-order support—then that may give us a reason to suspect that he is not as certain as he seemed about not needing surgery. To accept the second-order support shows some recognition of the fact that he may have the first-order need, surgery. This is the paradox; and, more importantly, it also works in the opposite direction. If John truly does not recognise his need for surgery, then he will not recognise that he has a reason to accept help to recognise that need. If a supporter says, 'we want to help you to see that you need surgery', then he is likely to reply 'but I don't need surgery'. Unless he is a particularly patient man, this might be followed by a comment about how nobody is listening to him.

To this use of the paradox, it might be objected that John could recognise his need for surgery yet refuse it. This raises two possibilities. An example of the first would be if John wished to live, and understood that he needed surgery to do so, but nevertheless decided that a religious prohibition made it impossible for him to accept the intervention. In this case, however, there is good reason to doubt whether John truly needs the surgery, no matter what his medical needs might be.[16] In contrast, an example of the second possibility would be if John wanted to live, and accepted that he needed surgery to live, but nevertheless refused the intervention without being able to give any reason why. In this case, however, what has occurred is not recognition. The concern in these situations is with a practical question, not with someone's ability to assent to theoretical propositions; so 'recognition' here necessarily includes the ability to make practical use of what has been recognised. This roughly accords with the natural use of the language. When, in this example, John agrees that he needs surgery to live, he has clearly assented to a proposition. After he refuses the surgery for no discernible reason, however, it would be unusual to nevertheless say that he 'recognises' his need for surgery. In conclusion, in the first example, John recognises an apparent need but denies that in the circumstances it is a need for him; and in the second what appears to be recognition is better characterised as mere assent. In neither case has John recognised a need for him and yet refused surgery.

Examining this paradox can help to reveal one condition necessary to justify systems, such as the MCA, that allow decisions to be made in someone's best interests against their will. For such decisions to be justified, it is not enough that someone fail to recognise a need; they must also fail to recognise that they have failed to recognise the need. If a person can recognise their own first-order needs, then they are capable of recognising their own best interests; so, in a liberal society, there is no obvious justification for giving someone else the power to make such decisions for them. Beyond

15 FL Moll and others, 'Management of Abdominal Aortic Aneurysms: Clinical Practice Guidelines of the European Society for Vascular Surgery' (2011) 41 European Journal of Vascular and Endovascular Surgery S1.

16 This point is dealt with in detail in the next section, but for now it is enough to observe that someone's objective needs cannot be identified without taking into account their own understanding of their life.

this, however, if someone does not recognise a first-order need but knows that they do not do so, then there is still no clear justification for giving someone else the ability to make decisions for them *against their will*. If, for example, John says 'I must be missing something, you decide for me', then the decision is not against his will. If, in contrast, he says 'I must be missing something, please explain it to me again', then there is no clear justification for ignoring his request for support and simply making the decision. When someone recognises their second-order need for help to identify their first-order needs, then less intrusive measures than a decision against their will are necessarily available. In other words, if it is assumed that self-determination is generally good for people and that any legal suspension of it should be for a reason, then the mere fact that people sometimes have needs that they cannot recognise is not enough to justify best interests decisions against someone's will. If such decisions are justified, it must be because sometimes people fail to make the second-order recognition that they cannot recognise a first-order need. That is to say, systems that allow best interests decisions against a person's will are only justified if people are sometimes caught in the paradox.

The argument in the previous paragraph does not demonstrate that best interests decisions against a person's will are ultimately justified. It may be that even if someone is caught in the paradox, then making a decision against their will is still unjustified for other reasons. For example, it might always cause significantly more harm to the person than respecting their preferences would. For this reason, reflecting on the paradox cannot settle the clash between the MCA and the UNCRPD discussed in the introduction. Nevertheless, if a second-order failure of recognition is necessary for a best interests decision against someone's will to be justified in a liberal society, then there are implications. In particular, the power to make such decisions will plausibly carry with it duties.

III. THE DUTY OF IDENTIFICATION

The previous section suggests that if best interests decisions made against someone's will are to be justified in a liberal society, then it is not enough that the person does not recognise a first-order need; they must also fail to recognise that they do not recognise a first-order need. From this general form of the argument, a specific form can be derived: in any particular case, a best interests decision against a person's will is only justified if they both fail to recognise a first-order need and fail to recognise that they are doing so. If it is assumed that decisions should only be made when it is justified to do so, then from this specific form of the argument a duty can be derived: the duty to identify whether someone is making such a second-order failure of recognition before deciding against their will.

There is, however, a problem; for this duty can become ensnared in the paradox that it is derived from. If the duty is fulfilled, a decision is only made when the person cannot be helped to make the second-order recognition that they have failed to recognise a first-order need. If, though, the person, such as John in the original example, is in this position, then they will deny that this is an accurate description of their situation. In other words, the same conditions that are necessary to justify the power to make a decision also make the use of that power unlikely to be accepted. This has profound consequences. So far, this article has largely taken a need as a simple objective

fact, but, of course, no supporter will ever have unmediated access to the objective facts. They will have, at best, a tangle of evidence from various sources about what someone's needs are. Furthermore, even a person's objective needs will be influenced by subjective factors. Even in the example, it was necessary to say that John wanted to live. A person's experiences and values partially shape their needs. This creates an evidential problem, for each of us has an access to these subjective elements that an assessor cannot equal. Atkins puts the point well: 'appreciation of the subjective character of experience brings with it the necessity for an epistemological humility' on the part of others.[17] If the supported person has better access to some of the evidence, then there is always the possibility that what appears to be a case of them failing to recognise a need is actually a case of third parties attributing needs to them that they do not have.[18]

Subjective elements shape our needs, so it will often be difficult to show that someone has needs that they are failing to recognise. This does not, however, mean that providing sufficient evidence is always impossible. Unfortunately, the UN Committee on the Rights of Persons with Disabilities appears to have concluded otherwise. It says that it is a flaw to presume to be able to 'accurately assess the inner-workings of the human mind'.[19] If this is a flaw it is, as Dawson points out, one that can be found throughout the legal system;[20] but, beyond that, the Committee fails to apply its own position consistently. Some evidence is only accessible to the supported person, but this is seldom the only relevant evidence that is only accessible to one person. Imagine John says that he doesn't need surgery, and supports his claim by saying that the doctors have kidnapped him in order to harvest his organs. In this case the staff, by virtue of *their* subjective experience, will have access to evidence that is not available to John. Some of this evidence, for instance, the doctor's intention to save John's life or her memories of learning how to interpret the relevant scans, will bear on the question of what John's needs actually are. The existence of subjective perspectives does not, by itself, call into question the existence of objective needs. It just means that these situations are evidentially complex.

Evidentially complex situations are not unusual. As Wimsatt points out, 'for the complex systems encountered in evolutionary biology and the social sciences, it is often unclear what is fundamental or trustworthy'.[21] A common response is to seek robust results: if different types of measurement all return the same result, then that result can be treated as more trustworthy.[22] This idea of seeking robust results is relevant here; for if the person concerned has access to evidence that cannot be gained in any other way, then this implies that their own assessment of the situation should

17 Kim Atkins, 'Autonomy and the Subjective Character of Experience' (2000) 17 Journal of Applied Philosophy 71.

18 Judges in the Court of Protection are aware of this danger. See, for example, *V v Associated Newspapers* [2016] EWCOP 21 [67]–[68] (Charles J).

19 General Comment (n 4) para 15.

20 Dawson (n 5) 74.

21 William C Wimsatt, *Re-engineering Philosophy for Limited Beings* (Harvard 2007) 56.

22 ibid ch 4. This process is not infallible. 'Robust detection' will break down if apparently independent measurements are not truly independent. Brett Calcott, 'Wimsatt and the Robustness Family' (2011) 26 Biology and Philosophy 284.

never simply be ignored. This in turn suggests that the duty of identification is only fulfilled if the person's own perspective on their needs is sought out and treated as direct evidence of whatever it is that they assert. Only if the countervailing evidence is overwhelming should the person's view be treated as evidence that they cannot recognise an objectively existing need. In practice, this may equate to Banner and Szmukler's application of Davidson's Principle of Charity: 'we must assume that a speaker is by and large consistent and correct in his beliefs'.[23] Indeed, if a supporter takes the opposite approach, and *begins* by taking a person's disagreement as evidence that they cannot identify their needs, then it is not the person who does not recognise that they have failed to recognise something. It is the supporter.[24] The approach suggested here will not, of course, resolve every possible evidential problem. It cannot address the question of what should first trigger the suspicion that someone does not recognise their own needs. It can only suggest what should be done when that suspicion is already present. Similarly, it cannot help if it is suspected that someone cannot recognise a need, but, despite every available support, they cannot or will not share their own perspective. This last point can, however, be accommodated by acknowledging that the duty of identification requires starting with the person's own perspective only when it is possible to do so.

If systems which allow best interests decisions against a person's will should respect the duty of identification, and the duty of identification requires someone's testimony to be treated as evidence of what they say before it is treated as evidence that they do not understand, then this offers two criteria by which the MCA can be assessed. The first criterion is whether it recognises the duty of identification at all, and the second is whether it recognises that the duty entails this careful treatment of the person's own testimony. It meets the first criterion. Capacity assessments under the MCA are of someone's ability to make a particular decision;[25] and an inability to make a decision is specified as an inability to understand, retain, use, or weigh relevant information; or an inability to communicate the decision.[26] The statutory language is wider than that derived from the paradox, but 'relevant information' includes the consequences of the decision,[27] and this seems close to being unable to identify your needs in that particular situation. This impression is reinforced when the Act's principles are examined. Section 1(2) requires that someone is assumed to have capacity unless the contrary is established. If being unable to understand or use information is analogous with the content of the duty of identification, then it is this section of the Act that makes it a duty to undertake a process of identification in every case. Furthermore, this principle is reinforced elsewhere in the Act. Section 5(1) requires that, for acts 'in connection with' care or treatment, 'reasonable steps' are taken to establish whether the person has capacity; and section 1(3) requires 'all practicable steps' be taken to

23 Donald Davidson, 'Psychology as Philosophy' in Donald Davidson (ed), *Essays on Actions and Events* (OUP 2001) 229, 238; cited in Natalie F Banner and George Szmukler, '"Radical Interpretation" and the Assessment of Decision-Making Capacity' (2013) 30 Journal of Applied Philosophy 379.

24 Thanks to David Gibson for this point.

25 MCA (n 3) s 2(1); *PC* v *City of York Council* [2013] EWCA Civ 478, [2014] 2 WLR 1 [35] (McFarlane LJ).

26 MCA (n 3) s 3(1).

27 ibid s 3(4).

help someone to make a decision before they are treated as unable to do so. This last section should distinguish between those who can be helped to recognise their needs and those who cannot, and only allow decisions in the latter case. The Act, especially this provision, has 'not been widely implemented';[28] but it certainly recognises the existence of something broadly similar to the duty of identification. Unfortunately, the same cannot be said of the second criterion, the careful treatment of the person's own testimony.

The Act's near silence on the question of evidence prevents it from requiring that the person's own testimony be treated with the care that the duty of identification demands. It does contain a few evidential prohibitions: incapacity cannot be established merely because of an unwise decision; or because of a person's condition, behaviour, age, or appearance.[29] It does not, however, indicate how the person's own perspective should be treated. The Code of Practice examines capacity assessments in more detail than the Act,[30] but it does not mention the person's perspective in this context. Although it discusses the possibility of challenging a capacity assessment,[31] the standard against which the assessment will be measured is that of the Act and Code;[32] so this takes things no further. Worse yet, an ambiguity in the language of the Code may mislead. It says that 'nobody can be forced to undergo an assessment of capacity'.[33] At face value, this would appear to mean that people have the ability to veto their own capacity assessment; although framed in an unhelpfully confrontational way, this would mean that their perspective would have to be taken into account. The Code has not been read in this way, as the '*DD*' series of cases illustrate. In these cases, there was consensus that 'threats or attempts to force *DD* to agree to an assessment are not acceptable',[34] but it was nevertheless declared lawful to remove her from her home 'by force' to a place where assessment could take place.[35] Her lack of cooperation with the subsequent assessment was respected;[36] but, importantly, this did not prevent her from being found to lack capacity to make decisions about litigation, contraception, and sterilisation.[37] In other words, *DD* could decline to take part in her capacity assessment, but that assessment took place without her anyway. There is no right to veto your own capacity assessment, and the Act contains no less extreme way to ensure that the person's own testimony is given a fair hearing, so the MCA fails to require that the second criterion of the duty of identification is fulfilled.

28 Select Committee on the Mental Capacity Act 2005, 'Mental Capacity Act 2005: Post-Legislative Scrutiny' (2013–14, HL Paper 139) para 79–83, 103–10.

29 MCA (n 3) s 1(4), 2(3). The word 'merely' is significant—these factors can have evidential force if they do not individually determine the question—*D* v *R (the Deputy of S)* [2010] EWCOP 2405, [2011] WTLR 449 [40] (Henderson J).

30 Department for Constitutional Affairs, *The Mental Capacity Act 2005 Code of Practice* (The Stationery Office 2007) ch 4.

31 ibid para 4.63–4.65.

32 ibid para 4.64.

33 ibid para 4.59.

34 *The Mental Health Trust v DD (No. 3)* [2014] EWCOP 13, 142 BMLR 156 [36] (Cobb J).

35 ibid [37]–[41].

36 *The Mental Health Trust v DD (No.4)* [2014] EWCOP 44 [5] and, [8]; The Mental Health Trust v DD (No. 5) [2015] EWCOP 4 [135] (Cobb J).

37 ibid *The Mental Health Trust v DD (No. 4)* [13]; *The Mental Health Trust v DD (No. 5)* [63], [80].

It seems likely that the Act's failure to ensure that someone's own testimony is treated as evidence contributes to such testimony being discounted in practice. Empirical studies show that a disagreement between an assessor and the person assessed can sometimes be taken as evidence that the person 'lacks insight', without first taking the disagreement as a reason to question the validity of the assessment.[38] A lack of insight appears to be strongly correlated with a finding of incapacity,[39] so the uncritical use of this concept is likely to have consequences. Professionals using 'insight' uncritically have been corrected by the Court of Protection,[40] but judges, too, sometimes seem to use the concept in an indiscriminate way.[41] When this happens, a person's disagreement can be discounted if they lack insight, and their disagreement is evidence that they lack insight;[42] so it becomes 'impossible to disbelieve a doctor and retain capacity'.[43] Here, the duty of identification has become entangled in the paradox that it was derived from; and it is the assessor, not the person assessed, who has failed to recognise what it is that they do not recognise. The Act allows this situation to be avoided. It does not, however, require that it be avoided.

IV. ADVOCACY AND THE DUTY OF IDENTIFICATION

Since the MCA was passed, the perspective of the decided-for person has become a more prominent issue; and the effects of this new prominence can already be seen in domestic law, particularly in changes to the legal framework governing the provision of advocacy. These developments bring the law closer to requiring compliance with the second criterion of the duty of identification, but there are, however, limits to what advocacy can achieve in this respect. This section discusses the recent legislation, then its limits. First, however, it examines the advocacy provisions of the MCA.

The MCA requires an Independent Mental Capacity Advocate (IMCA) to be appointed when someone is thought to lack capacity, and has no-one else to speak on their behalf; and faces serious medical treatment, a long-term change of accommodation, or a deprivation of liberty.[44] IMCAs can challenge capacity decisions,[45] so they offer an avenue for challenging assessments that ignore the person's own testimony. Nevertheless, unless appointed under section 39D of the Act when someone is deprived of their liberty,[46] they are consulted to help determine what would be in the

38 Val Williams and others, 'Making Best Interests Decisions: People and Processes' (Mental Health Foundation 2012) <http://www.mentalhealth.org.uk/content/assets/PDF/publications/best_interests_report_FINAL1.pdf?view=Standard> accessed 21 June 2015, 55–58; Charlotte Emmett, 'Homeward Bound or Bound for a Home? Assessing the Capacity of Dementia Patients to Make Decisions about Hospital Discharge: Comparing Practice with Legal Standards' (2013) 36 International Journal of Law and Psychiatry 73, 77.

39 GS Owen and others, 'Mental capacity, diagnosis and insight in psychiatric in-patients: a cross-sectional study' (2008) 39 Psychological Medicine 1389.

40 *KK* v *STCC* [2012] EWCOP 2136 [36],[64](Baker J).

41 *Islington LBC* v *QR* [2014] EWCOP 26, (2014) 17 CCL Rep 344 [96] (Batten DJ); *An English Local Authority* v *SW* [2014] EWCOP 43, [2015] COPLR 29 [20] (Moylan J).

42 Kate Diesfeld and Stefan Sjöström, 'Interpretive Flexibility: Why Doesn't Insight Incite Controversy in Mental Health Law?' (2007) 25 Behavioural Sciences and the Law 85.

43 Neil Allen, 'Is Capacity "In Sight"?'(2009) 19 Journal of Mental Health Law 165, 167.

44 MCA (n 3) ss 37–39E.

45 The Mental Capacity Act 2005 (IMCAs) (General) Regulations 2006 reg 7(1)(b).

46 *AJ v A Local Authority* [2015] EWCOP 5, (2015) 18 CCLR 158 [107–11].

person's best interests.[47] This leads to an ambiguity. If the IMCA believes that the person lacks capacity, then they might not challenge an assessment that the person disagrees with; for they might not believe that it is in the person's best interests to do so.[48] In other words, IMCAs are merely allowed, not required, to argue for the person's own perspective.

The advocacy provisions in the Care Act 2014 are more robust than those in the MCA, and provide a useful contrast, as the trigger for advocacy is still framed in terms of ability to make a decision. The newer Act requires an advocate to be appointed during needs assessments, care planning, or safeguarding processes;[49] if 'the individual would experience substantial difficulty' making a decision, and no-one else can represent and support them.[50] This is a lower bar than in the MCA, for 'substantial difficulty' does not require the person to be thought to lack capacity;[51] but a more relevant difference between the two Acts is in what, exactly, the advocate is required to do once appointed. Under the Care Act, an advocate is *required* to communicate the 'views, wishes, or feelings' of a person thought to lack capacity,[52] and the statutory guidance reinforces this as a duty to put forward the person's own case.[53] This goes further than the MCA towards robustly fulfilling the second criterion of the duty of identification, and it may also influence the older Act. The Law Commission has suggested that IMCAs 'be replaced by a system of Care Act advocacy'; this would incorporate the more robust standard into the MCA itself.[54]

Strengthening the law governing the provision of advocacy is a step towards requiring compliance with the duty of identification, but it has limits. These limits are of three types: limits of scope, limits in implementation, and cultural limits. The legislation itself has limited scope. Only certain decisions trigger the duty to appoint an advocate. Between the two Acts, many major decisions will be covered, but it is not clear that all will be; and, as has been acknowledged in Parliament, 'minor' decisions can be as important to the person concerned as 'major' ones.[55] In practice, however, limitations in implementation are likely to be more severe than those of scope. Failures of implementation were the most consistent theme of the recent House of Lords Select Committee Report into the MCA,[56] and IMCAs are not always appointed even when it is legally required.[57] Legislation requiring advocacy of a certain level is not, by itself, sufficient to ensure that practice will consistently reach that level;

47 MCA (n 3) ss 37(1)(b), 38(1), 39(1)(b), 39A(1)(b), 39C(1)(c).
48 Select Committee (n 28) para 167.
49 Care Act 2014 ss 67(1), (3), 68(1).
50 ibid ss 67(4) and 68(3).
51 Department of Health, 'Care and Support Statutory Guidance' (2014) paras 7.55, 7.58.
52 The Care and Support (Independent Advocacy Support) Regulations 2014 reg 5(7).
53 Statutory Guidance (n 51) para 7.43.
54 Law Commission, 'Mental Capacity and Deprivation of Liberty: A Consultation Paper' (Consultation Paper No 222, 2015) para 9.41.
55 Joint Committee on the Draft Mental Incapacity Bill, 'First Report'(2002–2003, HL Paper 189-I, HC 1083-I) para 27.
56 Select Committee (n 28).
57 Department of Health, 'The Seventh Year of the Independent Mental Capacity Advocacy (IMCA) Service' (2015) <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/416341/imca-report.pdf> accessed 10 August 2015, 34.

and the current political and economic climate is relevant here. Local authority budgets have fallen by almost 40% in real terms over the last five years, and authorities have lost their ability to insulate social care from the effects of this drop in funding; yet the overall need for care is increasing dramatically.[58] In these circumstances, advocacy may be neglected; and the provision of more 'basic' physical care prioritised.

Scope and implementation offer relatively contingent limits to advocacy. Changes to the legislation, or to practice and funding, could overcome them. There are, however, also cultural limits to the ability of an advocate to fulfil the duty of identification, and these will be harder to overcome. These cultural limits gain their force from the paradox, but to show this requires careful disambiguation. The example of John can help. If an advocate wants to represent John's own perspective as faithfully as possible, then she faces a choice. John both wishes to live and wishes to refuse life-saving surgery. If she presents either of these wishes alone, then she will not fully represent John's relevant views, but if she presents both, then she will undermine John's case by emphasising its contradictions. In these cases, then, presenting what someone wants without explaining why they want it is insufficient. Instead, an advocate should present the person's overall view, for example, John's view that refusing surgery will not risk his life, and the *reasons why* that view should prevail.[59]

If an advocate attempts to present the reasons why someone's view should prevail, then another choice reveals itself. This is the choice between only presenting John's reasons for his view and also presenting other reasons for reaching the same view. With regards to the duty of identification, these are different things. If the advocate presents John's reasons for his view, then she is helping to ensure that his perspective is treated as evidence, as the duty requires. If, however, she offers reasons which are not John's for why John's wishes should prevail, then a shift in argumentation has occurred. Proposing reasons that are not John's does not help to show that John understands the relevant information and hence has capacity. Instead, the advocate is effectively arguing that regardless of whether John has capacity or not, his wishes should prevail. This is no longer about the duty of identification; it does not bear on the question of whether John recognises his needs. There are good reasons for keeping these two lines of argumentation distinct. Being found to have capacity is not the same as being found to lack capacity but nevertheless having your wishes followed. A finding of incapacity can carry costs to the person beyond simply not getting what they want: for instance, self-stigmatisation and loss of confidence.[60] Beyond this, it is only by disambiguating between the two that the cultural limits on advocacy can be made clear.

The distinction between advancing someone's own reasons for their view and presenting other reasons for why that view should prevail makes it easier to see the cultural limits on the ability of an advocate to fulfil the duty of identification. An advocate's position will often, regrettably, give her more credibility than the person themselves, but this credibility has limits, and some of these limits are cultural. If,

58 National Audit Office, 'Care Act First-Phase Reforms' (2015–2016, HC 82) paras 1.6–1.9.

59 The same point is true of 'wishes and feelings' in best interests decisions: Laura Pritchard-Jones, 'This Man with Dementia – Othering the Person with Dementia in the Court of Protection', Medical Law Review (forthcoming).

60 Bruce J Winick, 'The Side Effects of Incompetency Labelling and the Implications for Mental Health Law' (1995) 1 Psychology, Public Policy, and Law 6.

again, John is paranoid and believes that he has been kidnapped so that his organs can be harvested, but is actually ill in hospital, then an advocate arguing that his reasons for refusing treatment are literally true will not help him. She will only harm her own credibility. For a reason to be convincing, it must have some coherence with the account of the world that listeners assume to be true. This is not, by itself, a problematic limit: in John's case, even a charitable treatment of his testimony should almost certainly conclude that he is failing to recognise a need. It does, however, indicate where the problematic cultural limits to advocacy are. Experimental psychology has shown the existence of implicit social biases. These biases are entirely distinct to someone's explicit attitudes, they are robust, and they predict behaviour.[61] In other words, people consistently act on biases that they deny, and may not even perceive. In particular, strong implicit biases against those with disabilities have been shown to be widespread, even among professionals who work closely with them.[62] These biases, combined with the need for a person's reasons to have some minimal coherence with commonly held assumptions to be credible, are sufficient to impose limits on advocacy's power to fulfil the duty of identification. Recognising the existence of widespread unacknowledged bias entails recognising the risk that sometimes when a person's reasons seem unconvincing, because they do not cohere with the commonly accepted view of the world, it is society, not the person, which has not recognised its failure to recognise significant facts. Furthermore, when the paradox reappears in this societal form, it limits advocates in two ways. First, advocates may be implicitly biased; for explicit opposition to discrimination does not guarantee that someone is not implicitly biased.[63] Secondly, if the advocate is not biased, and others are, then arguing for the person's perspective will risk the advocate's credibility in the same way, if seldom to the same extent, that arguing John really was going to have his organs harvested did. The advocate will literally be arguing against what 'everyone knows'. These cultural limits will be difficult to address: implicit biases may be difficult to change,[64] and such biases against those with disabilities do not seem to have fallen in the last decade.[65] They are, however, only limits. An area of freedom should remain in which an advocate can challenge biases without destroying their own credibility; and if that balance is difficult to find, then this is an argument for strong advocacy, not against it. Nevertheless, the cultural limits on advocacy, in combination with the limits of scope and in implementation, do mean that while stronger advocacy legislation is a step towards requiring that a person's testimony is treated with due care, it is only one step. Others will almost certainly be needed.

61 Kristin A Lane and others, 'Implicit Social Cognition and Law' (2007) 3 Annual Review of Law and Social Science 427; Brian A Nosek and others, 'Implicit Social Cognition: From Measures to Mechanisms' (2011) 15 Trends in Cognitive Sciences 152.

62 Michelle Clare Wilson and Katrina Scior, 'Attitudes Towards Individuals with Disabilities as Measured by the Implicit Association Test: A literature review' (2014) 35 Research in Developmental Disabilities 294.

63 Lane and others (n 61) 431.

64 Jennifer A Joy-Gaba and Brian A Nosek, 'The Surprisingly Limited Malleability of Implicit Racial Evaluations' (2010) 41 Social Psychology 137.

65 Wilson and Scior (n 62) 319.

V. THE DUTY TO SUPPORT

In addition to the duty of identification, a duty to provide support can be derived from the paradox; and, as with the first duty, this offers a criterion by which legal systems that permit best interests decisions against a person's will can be evaluated. At the outset, however, it should be stressed that 'support' in this context is limited to support of one particular, relatively narrow, type. It is support to someone who has had a decision made for them because of a second-order failure of recognition, which is directed at helping them to avoid such failures in the future. There are further limits on this support—in particular, it cannot be against the person's will—that are discussed below. First, however, this section shows how the duty follows from the paradox.

As argued in Sections II and III, deciding against a person's will can only be justified if someone cannot recognise that they do not recognise a need. At the same time, however, interfering with someone's self-determination is understood in liberal societies to be wrongful. Indeed, this understanding motivates the general principles of both the MCA[66] and the UNCRPD.[67] On a simplistic reading, this general prohibition on interfering with someone's self-determination is entirely displaced when their second-order failure of recognition justifies making a decision against their will. Such a reading is, however, *too* simplistic; for it collapses the difference between doing an unqualifiedly right act and doing a justified wrong.[68] Making a decision against someone's will because of their second-order failure of recognition is not an unqualifiedly right act. The very structure of this description denies it. It presents an apparent wrong, making the decision against their will; and then presents a reason, the failure of recognition, for nevertheless doing it. It is, at best, a justified wrong; and this is important. Justified wrongs are necessarily cases in which moral conflicts have occurred and been resolved; but, as Williams observes, such conflicts are not 'all soluble without remainder'.[69] One 'remainder' he examines is the regret that an 'admirable moral agent' might feel, even when they are acting as well as circumstances seem to allow.[70] This is a good starting place; for it seems right to feel some regret if we have made a decision against someone's will, even if we are sure that doing so is justified in the particular case. Furthermore, it seems right that, as Williams says,[71] someone experiencing such regret might conclude that they ought to avoid such situations arising in the future. In the particular case, the moral conflict has arisen because someone has not recognised that they do not recognise a need; so a regretful decision-maker might conclude that, where possible, they should help the decided-for person to avoid such second-order failures of recognition in the future. This, then, is the derivation of the duty to support. If it is accepted that a decision against a person's will is not an unqualifiedly good act but a justified wrong, that the appropriate response to inflicting

66 MCA (n 3) s 1.
67 UNCRPD (n 1) art 3(a).
68 Rosalind Hursthouse, *On Virtue Ethics* (OUP 2001) 44–49 discusses this distinction.
69 Bernard Williams, 'Ethical Consistency' in Geoffrey Sayre-McCord (ed), *Essays on Moral Realism* (Cornell UP 1988) 41, 52.
70 ibid 49.
71 ibid 50.

this justified wrong is regret, and that the appropriate response to such regret is to attempt to avoid such situations occurring in the future, then there is an apparent duty once a decision has been made against someone's will to attempt to help them to avoid second-order failures of recognition in the future. This is an endoxic derivation, it appeals to widely held evaluative beliefs;[72] so it can be rejected by someone who denies regret is an appropriate response to inflicting a justified wrong, or denies that we should act on such regret. Nevertheless, neither of those positions seems particularly appealing, so it does have some force. The duty to support, however, also faces an internal challenge.

Just like the duty of identification, the duty to support can become ensnared by the paradox that it is derived from. If the justification for making a decision against someone's will is that the person could not be helped to recognise their failure to recognise a need, then to claim that there is, nevertheless, a duty to provide this same help seems absurd, doomed to failure from the outset. This objection may hold true in some cases; and in those cases, if they can be reliably identified, there is no obvious duty to support. Few cases, however, are likely to be of this type. Just because we cannot help someone to recognise their own failure of recognition at one time does not mean that we cannot do so over a longer time period. Therefore, despite the entanglement of the paradox, the duty to support will still exist in most cases. It will, however, be limited in another way; for it cannot be used to justify decisions against a person's will. Before discussing this limit, however, it is worth evaluating whether the MCA recognises the duty at all.

The MCA may seem to recognise the duty to support. After all, section 1(3) requires 'all practicable steps' be taken to support someone to make a decision; and section 4(4) that someone is encouraged and permitted to improve their ability to take part in decisions affecting them, insofar as that is 'reasonably practicable'. These provisions do not, however, require fulfilment of the duty to support; and this is, in part, due to the same 'decision-specific' structure of the Act that buttresses the duty of identification. Section 1(3) only requires that someone be helped to make the particular decision in question. It does not create any duty that continues after the decision is made, but it is then that the duty of support arises. Similarly, section 4(4) requires that the person's 'ability to participate' be supported during best interests decisions, but it, too, is specific to the particular decision being made. It does not create any ongoing duty.[73] When a cohesive team is working well with the person, then an ongoing duty to support might, nevertheless, emerge from these sections. As Series points out, however, both decision-making and support under the MCA are 'dispersed over a large number of disparate actors',[74] so the Act does nothing to encourage this sort of practice. Therefore, as with the duty of identification, the MCA allows the duty to support to be recognised, but does not require it.

72 Martha C Nussbaum, 'Aristotle, Nature, and Ethics' in JEJ Altham and Ross Harrison (eds), *World, Mind, and Ethics: Essays on the Ethical Philosophy of Bernard Williams* (CUP 1995) 86, 100.

73 Indeed, it only arises when 'the time for supported decision-making is past'. *Re NRA* [2015] EWCOP 59 [56] (Charles J).

74 Series (n 5) 84.

It may be that the MCA is the wrong place to look for the duty to support. The provision of support, beyond that required for the immediate decision, could simply be beyond its remit. After all, as Lady Hale says, the Act is 'concerned with enabling the court to do for the patient what he could do for himself if of full capacity, but it goes no further'.[75] This point has been reinforced in the Court of Appeal: acting in someone's best interests grants no additional power 'to obtain resources or facilities from a third party'.[76] Given this, it is worth again looking to the Care Act 2014. After all, it gives local authorities duties to promote 'control by the individual over day-to-day life'[77] and to 'reduce the needs for care and support of adults in its area'.[78] At first sight, it seems possible to read a duty to help someone to recognise their own needs into these provisions, but, unfortunately, the detail of the Act makes this unlikely to consistently happen. Local authorities only have a duty to provide for needs that meet the 'eligibility criteria',[79] and these criteria refer only to first-order needs, such as maintaining nutrition.[80] Some service providers may decide that supporting someone to identify their own needs in the long term is the best way to help them to meet those needs; and local authorities can provide services when they have no duty to do so.[81] Nevertheless, as with the MCA, there is nothing in the Care Act that makes fulfilling the duty to support required, instead of merely allowed.

Beyond failing to require the duty to support, the MCA also fails to recognise the limits of support. Words like 'support' can be slippery; and it is easy to envisage a duty to help someone to recognise that they have not recognised their needs quietly becoming a power to force someone to accept our assessment of their needs. Berlin's famous warning, 'to manipulate men, to propel them towards goals which you – the social reformer – see, but they may not, is to deny their human essence',[82] is as apt here as it is anywhere. One way to avoid this slippage is to return to the paradox, and to remember that deciding against a person's will is only justified when they do not recognise their failure to recognise a first-order need. If making a decision against someone's will does not make it possible to satisfy the first-order need, then this justification fails. This is what happens with the duty to support. If we say 'due to the paradox, John cannot be helped to recognise his need for surgery, so we must decide for him', then a decision to operate can be implemented without any further input from John. In contrast, if we say 'John cannot be helped to recognise his need to recognise his needs, so we must decide for him', then we cannot make the second-order recognition happen without John's input, no matter what we decide. The subjective character

75 *Aintree University Hospitals NHS Foundation Trust* v *James* [2013] UKSC 67 [18].

76 *Re MN* [2015] EWCA Civ 411 [80] (Munby P). This case has been granted permission to appeal to the Supreme Court.

77 Care Act (n 49) s 1(2)(d).

78 ibid s 2(1)(c).

79 ibid s 18(1).

80 The Care and Support (Eligibility Criteria) Regulations 2014, SI 2015/313, reg 2(2) The regulations do mention 'accessing and engaging in work, training, education or volunteering' —reg 2(2)(h)—but this refers to someone's ability to access existing programs. There is nothing to indicate that it requires the provision of the support discussed here.

81 Care Act (n 49) s 19(1).

82 Isaiah Berlin, 'Two Concept of Liberty' in Henry Hardy and Roger Hausheer (eds), *The Proper Study of Mankind: An Anthology of Essays* (Pimlico 1998) 191, 203.

of recognition dictates that it can only be reached by minimally consensual means; so the duty to support cannot justifiably be extended to decisions against a person's will. The MCA, however, does not distinguish between these two different situations; so it permits the meaning of 'support' to slip beyond reasonable bounds. For instance, in *Northamptonshire Healthcare NHS Foundation Trust* v *ML,*[83] a young man, 'ML', was given an 'opportunity to fulfil his potential' to be independent. The 'opportunity' involved a detention of up to two years in hospital and involuntary treatment that he would undoubtedly find traumatic.[84] It is unlikely that he experienced this as an opportunity or as support; and if he did not, then it must be doubted whether it really was these things. The MCA both fails to require respect for the duty of support, and fails to prevent 'support' sliding into unjustified coercion.

VI. CONCLUSION

Meno's paradox clarifies one condition necessary for the justification of a best interest decision against someone's will. It suggests that it is not enough that someone has failed to recognise a first-order need, but they must also fail to recognise their own failure of recognition. From this necessary condition, two duties can be derived: a duty to identify that any particular person is in this position before deciding against their will; and a duty to support those decided for to avoid this situation in the future. Any particular legislation that allows best interests decisions against a person's will can be measured against these duties. The MCA allows the fulfilment of these duties, but it does not robustly require them. Given this, it is perhaps unsurprising that the recent House of Lords report found the 'empowering ethos' of the Act's general principles has yet to be realised.[85] The failure to require compliance with the two duties is not due to those principles. It is, rather, due to the finer details of the Act and other legislation; some of which could, perhaps, be changed. The duty of identification requires that, during capacity assessments, the person's testimony is treated primarily as evidence of what they assert and only secondarily as evidence that they cannot recognise their needs. Changes to the legal framework governing the provision of advocacy has already brought the law closer to doing this, but there are important limits to what advocacy can achieve by itself. The second duty, the duty to support, could be given a legal basis by making a person's inability to recognise their other needs an eligible need under the Care Act. The MCA, however, also allows 'support' to slip into unjustified coercion, and this problem would have to be addressed separately.

This article shows that the MCA only allows, and does not robustly require, respect for two duties that any liberal system that allows best interests decisions against a person's will should be committed to. It does not, however, demonstrate whether or not best interests decisions are ultimately justified; and, given this, it does not support either the maintenance of the status quo, or the abandonment of all best interests decision-making that has been called for under some interpretations on the UNCRPD. Nevertheless, Meno's paradox also presents a challenge to those who favour the latter view; for they would appear to be committed to one of three following responses to

83 [2014] EWCOP 2, [2014] COPLR 439 [45] (Hayden J).
84 ibid [24], [59].
85 Select Committee (n 28) para 104.

it: either that no one is ever caught in the paradox; or that those who are so caught should be allowed to suffer whatever avoidable harms follow from their second-order failures of recognition; or that third parties can interpret someone's 'will and preferences' in ways that contradict what the person actually says. A detailed examination of these positions is beyond the remit of this article, but the first response appears false, the second callous, and the third sinister. Despite this, taking seriously the response to the paradox discussed here would seem to narrow, without entirely closing, the gap between the MCA and the UNCRPD. It would require increased attention to the person's perspective during capacity assessments, and at least one form of additional, consensual support. These are both things that align well with the overall aims and ethos of the UN Convention.

International Journal of Law and Psychiatry 53 (2017) 102–110

Contents lists available at ScienceDirect

International Journal of Law and Psychiatry

Humility when responding to the abuse of adults with mental disabilities☆,☆☆,★

Paul Skowron *

Centre for Social Ethics and Policy, University of Manchester, Manchester, UK

ARTICLE INFO

Article history:
Received 25 November 2016
Received in revised form 21 March 2017
Accepted 23 April 2017
Available online 20 May 2017

Keywords:
Humility
Abuse
Autonomy
Protection
Mental disability
Mental capacity

ABSTRACT

Legal theorists often reduce the ethics of responding to the abuse of another person to a clash between the principles of autonomy and protection. This reduction is a problem. Responding to suspected abuse requires humility – the potential responder must be aware of and respect their own limits – but humility cannot be usefully reduced to protection and autonomy. Using examples from the Court of Protection of England and Wales, this article examines the different ways that someone responding to abuse should respect their own limits, and suggests that a failure to do so will disproportionately affect people with mental disabilities. It is therefore necessary to attend to whether the law fosters humility among those who respond to abuse, although this must be tempered by humility about legal reform itself. Finally, the article shows how attention to humility can assist the interpretation of Article 16 of the UN Convention on the Rights of Persons with Disabilities; and suggests that, so interpreted, the Convention may help to promote humility when responding to abuse.

1. Introduction

Legal theorists often reduce the ethics of responding to the abuse of another person to a clash between autonomy and protection.[1] This causes a problem, one of a type observed by Williams (1985/2011). Theory aims for systemic unity, so it reduces the number of ethical ideas to a manageable minimum. 'Reflective criticism', in contrast, aims for shared understanding, so it uses any material that 'makes some sense and commands some loyalty' (Williams: p. 129–130). In other words, a choice must usually be made between theoretical simplicity and doing justice to the complexities of human life. This article is directed to the latter end. It suggests that responding to suspected abuse requires humility, something that cannot be usefully reduced to protection or respect for autonomy. It makes this argument in a series of widening concentric circles. Starting with a potential responder's attitude to their own knowledge, it expands to humility in individual actions, then to humility and legislative actions, and finally to humility and the UN Convention on the Rights of Persons with Disabilities ('UNCRPD').[2] Throughout, the argument is supported with examples taken from the Court of Protection of England and Wales.

A lack of humility when responding to abuse can lead to systematic injustices for people with mental disabilities. It does not, however, do so in an obvious way. This is not a case of direct discrimination, or even of the discriminatory application of formally neutral laws, although both of those things also occur. Instead, it is a case where the background conditions of society mean that applying apparently neutral, but inadequate, systems in an apparently neutral way will nevertheless disproportionately affect a particular group. People with mental disabilities suffer more abuse than those without (Hughes et al., 2012). This is a primary systematic injustice. If responses to abuse, however, are inadequate, then this also leads to a secondary systematic injustice. Mentally disabled people, because they suffer disproportionate abuse, disproportionately rely on social and systemic responses to abuse. If those responses are inadequate, then the burden of that inadequacy will disproportionately fall on them. This secondary injustice could, of course, be addressed by stopping the abuse; but it could also be addressed by improving the flawed system. Given the probable difficulty of ending widespread abuse, it may be best to attempt both of these things.

Sections 3 and 4 show, by example, that the principles of autonomy and protection cannot adequately guide responses to abuse. The reason why they are inadequate can be easily summarised. Analysing a case of

☆ This paper forms part of a special edition on Article 16 of the United Nations Convention on the Rights of Persons with Disabilities. The full contents of this special edition may be found at http://www.journals.elsevier.com/international-journal-of-law-and-psychiatry/.

☆☆ Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

★ Conflicts of interest: None to declare.

* School of Law, Williamson Building 4.42, University of Manchester, Oxford Road, Manchester M13 9PL, UK.
E-mail address: paul.skowron@postgrad.manchester.ac.uk.

[1] This binary is ubiquitous, but for examples see Ellis (1992), Richardson (2007), and Dunn, Clare, and Holland (2008).

[2] Convention on the Rights of Persons with Disabilities (adopted 13 December 2006, entered into force 3 May 2008) 2515 UNTS 3.

http://dx.doi.org/10.1016/j.ijlp.2017.04.005

suspected abuse in terms of this binary focusses attention on the relationship between the potentially abused person and the potential abuser, but this oversimplifies the situation. For a response to be possible, then there will not only be a relationship between two people, abused and abuser; but a set of relationships between at least three: abused, abuser, and potential intervener. The particular features of a potential intervener, such as what they know and what they can do, cannot be ignored when deciding how they should respond to suspected abuse. If that is so, though, then any potential intervener should not only evaluate the situation in front of them. They should also evaluate their own relationship to that situation. They should, as Nagel says (1986: p. 4), 'step back from [their] initial view … and form a new conception which has that view and its relation to the world as its object'. This 'step back' is not protection or respect for autonomy. It is something different, a form of self-evaluation. It is not enough, however, to merely say that potential responders to abuse should evaluate themselves. That would provide no information about which parts of the self should be evaluated or to what standards it should be held. If the idea of self-evaluation is to be useful, more detail must be provided. That might seem like a massive task. Fortunately, though, there is no need to start from scratch. History provides a rich concept of self-evaluation that has been analysed for centuries: humility.

2. Humility

This section describes some features of humility.[3] Later ones use this description to analyse responses to abuse. One feature of humility, and enduring theme of the literature, is central to this article: humility includes an awareness of, and respect for, your own limits.[4] If humility is to augment the idea of self-evaluation, then this can provide a starting point. It suggests which part of themselves a potential responder to abuse should evaluate, their limits. Beyond this, however, humility can also help to provide the standards with which someone can evaluate themselves. This can be shown by analysing a succinct statement by Aquinas: 'Humility restrains the appetite from aiming at great things against right reason' (1247/1947: IIa–IIae, 161, 1).

Aquinas's statement has three important features. First, humility 'restrains the appetite', it leads someone to re-evaluate a course of action that they would otherwise be inclined to take. In the context of suspected abuse, this can work in either direction. Humility can moderate a desire to protect someone or a desire to respect their autonomy, when either is against 'right reason'. This is the second point. Humility does not restrain desire indiscriminately, it targets only desires that are unreasonable in the circumstances. Third, and finally, the ends that humility restrains someone from aiming at are 'great'. Humility is not important because it stops someone from trying to do wrong, other virtues do that, but because it stops them from unreasonably trying to do right. The desire to protect another person and the desire to respect their autonomy are both admirable. It is, however, possible to attempt either in an unreasonable way, as examples in later sections illustrate.

Aquinas's statement makes it possible to address some persistent misunderstandings about humility. It is important to do so. Humility does not have the prominence that it had in earlier eras. Consequently, modern scholars can fail to grasp not only the subtlety of older traditions (Clement, 2015: p. 15–20) but also the central role of human limitation in enlightenment philosophies (Cooper, 2013: p. 5), which are sometimes assumed to have consigned those traditions to history. This is not to say that humility is now entirely neglected, for it has been defended as essential to democracies (Button, 2005), especially pluralistic ones (Parens, 1995). Nevertheless, three misunderstandings are common: a conflation of humility and self-abasement; the idea that humility necessarily reinforces existing hierarchies; and the idea that humility is overly inwards and self-directed.

The first common misunderstanding is to equate humility with self-abasement, so to think that it must necessarily lead to inaction or despair. This view fails to notice that someone is only humble when they recognise a limit that they actually have. Being aware of real limits does not require someone to imagine limits that do not exist. When Murdoch (1970: p. 93) says that 'humility is not a peculiar habit of self-effacement, rather like having an inaudible voice, it is selfless respect for reality', her point is orthodox. The emphasis on right reason in Aquinas has already been mentioned; and, similarly, Augustine (415/1997: Ch.38) states that humility should be 'on the side of truth, not on the side of falsity'. Tying humility to honesty in this way undermines the idea that humility leads to inaction, and this is significant for responses to abuse, where anything that leads to unnecessary inaction might seem suspect. The connection to honesty also reveals something of the deeper rationale for humility. One reason that humility is important is that humans do systematically tend to overlook their own limits. This insight is old, but experimental psychology tends to support it (Dunning, 2005); and just as importantly, psychology also suggests that this overconfidence can sometimes be addressed (Ehrlinger, Mitchum, & Dweck, 2016). In other words, science seems to reaffirm both the need for humility and its possibility. This, too, is relevant to responses to abuse, where demanding the impossible would be little practical help.

The second common misunderstanding of humility is to think that it necessarily maintains hierarchies; that it is something forced on the weak by the strong. This misunderstanding, like the last, is deeply ahistorical. It does not pay attention to where the advice to be humble was traditionally directed. Although the virtue has been invoked to support existing social orders, there are also influential traditions that assert it is the powerful who most need humility if their power is to remain legitimate. Clement (2015: p. 127) shows that humility in the early modern period was widely understood as 'a virtue that resists tyranny even as it can be invoked to support the status quo'; and, similarly, Klancer (2012: p. 670) shows that for both Aquinas and the Confucian Zhu Xi, humility is 'a virtue inherently fitting for the strong'. The idea that it is the powerful who must be humble has more recently been applied to clinical relationships (DasGupta, 2008); and this aspect of the virtue is important to the present article, for those in a position to intervene in cases of abuse almost invariably have considerable power relative to the person thought to be suffering abuse.

The final common misunderstanding of humility is to think that because it is concerned with the person's own limits that it is too inwardly directed to usefully guide action. Once again, this idea lacks historic support. As Button (2005: p. 850) notes, even St. Bernard, a monk writing primarily for other monks, emphasised humility not as an end in itself, but as a precondition to cultivating good relationships. Indeed, Bernard (1120/1973: III.6) was blunt: 'you will never have real mercy for the failings of another until you know and realise that you have the same failings in your soul'. Similarly, Boyd (2014: p. 258) points out that for Aquinas, too, humility is social. It 'enables us to value ourselves as members of a community in which no one person possesses independent, god-like status'. The more recent literature has, if anything, further emphasised the social implications of humility: Button (2005) argues that pluralistic democracies depend on citizens cultivating sensitivity to their own limits, and the limits of their institutions and laws; Penrose (2010) that humility underwrites the correct moral response to wrongdoers; and Coulehan (2011: p. 206) that humility is necessary to respond appropriately to the 'ambiguities, mysteries, and surprises' of the clinical encounter. This article is a further small contribution to the social argument for humility. Its central claim is that only humble responses to abuse will be socially effective, both at the level of

[3] Humility has a long and complex history. The aim here is to draw attention to some relevant details, not to give a complete account of the virtue.

[4] Compare three statements almost evenly spaced through the last seven centuries: Aquinas (1274/1947: IIa-IIae, 161, 2) 'Knowledge of one's own deficiency belongs to humility'; Descartes (1649/1985: art.155) 'We have humility as a virtue … as a result of reflecting on the infirmity of our nature'; and Grenberg (2005: p.162) 'The humble person understands her limits'.

the abused person's relationships and at the level of the whole society. To show this, though, more detail is needed on the particular limits that a potential responder to abuse should recognise and respect. The next section introduces the first type of limit, limits to knowledge.

3. Humility about knowledge when responding to abuse

3.1. Humility about what is known

This is the first of two sections discussing humility when confronting a particular case of suspected abuse. It addresses humility about knowledge, or 'epistemic humility'. The need for this sort of humility is illustrated using a case from the Court of Protection of England and Wales, *Somerset County Council v MK ('Somerset')*.[5] This case concerned a young woman, 'P', with severe learning disabilities and autism who had lived with her family for her entire life.[6] Bruising was noticed on P's chest, and this quickly led to an investigation by social workers that concluded that it was 'highly likely that P has received a significant injury from someone or something other than herself'.[7] Consequently, she was removed from her family and transferred to a 'respite placement' and then to an 'assessment and treatment unit'.[8] Clearly, this action was intended to protect P, but it could also be plausibly presented as furthering her autonomy. Abusive relationships do not respect autonomy; and, beyond that, 'autonomy-impairing fear' may be a major reason that such relationships continue (Nedelsky, 2012: p. 312). If so, it can be plausibly argued that removing someone from an abusive relationship might promote their autonomy. In *Somerset*, then, protection and respect for autonomy could be easily made to work together to justify removing P from the family home. This is why epistemic humility was also needed. Just prior to the bruising being found, P had been witnessed hitting herself in the chest on a school trip and had been physically restrained by staff. These events would have been 'easily discoverable by the social workers if they had carried out a proper investigation'.[9] In the year that elapsed before these easily discoverable facts were recognised, P spent time deprived of her liberty in a placement that was 'clearly not appropriate'. This led to anti-psychotics being prescribed to manage her behaviour.[10]

The initial suspicion that P was suffering abuse was reasonable, but something clearly went wrong in the response to that suspicion. This wrongness cannot be adequately caught using only the concepts of protection and autonomy. The inadequacy of these two concepts, however, may not be obvious. It might be objected that removing P from her family did not protect her, it exposed her to new risks of emotional harm and inappropriate medication. Similarly, it did not respect her autonomy, it required overriding her consistently expressed wish to go home.[11] The social workers, according to this argument, failed to protect P or to respect her autonomy; so these concepts are sufficient to express what went wrong. This analysis, however, rests on a fundamental error. Criticising the social workers for failing to protect P or respect her autonomy assumes that they knew then what we know now. To assume this is, however, to miss the point in the most dramatic way possible. The social worker's culpable fault was that they *did not know* something that they should have known. If it was possible to travel back in time and talk to the professionals who interacted with P, then it would not help to tell them to protect her. All of the evidence indicates that they thought that they *were* protecting her. Similarly, it would do no good to tell them to respect her autonomy. The evidence also indicates that they thought that they *were* respecting her autonomy. Indeed, even after conceding that there were no 'safeguarding issues', the Local Authority continued to argue that P should not be returned to her family, so that she could better fulfil 'her wider potential'.[12] This, however misplaced, is an argument from autonomy. Just because hindsight shows that the Local Authority failed to protect P or to respect her autonomy does not mean that these concepts could have adequately guided decision-makers at the time.

If it was possible to travel back in time and talk to the professionals who interacted with P, then it might have helped to draw their attention to how little they really knew about the situation. Had the social workers paid more attention to the limits of their knowledge, then it is reasonable to believe that they would have investigated more fully. In addition to evaluating how to protect P and respect her autonomy, they also needed to evaluate themselves. Indeed, in court, the senior social worker admitted this, saying that there was 'a failure of those on the ground to realize that they are out of their depth'.[13] Furthermore, there is evidence that the wider 'systemic failure'[14] of the Local Authority was also, at least in part, a failure of epistemic humility. Actions were repeatedly taken without consulting those who knew P best,[15] and P's own attempts to communicate her wish to go home were perversely misinterpreted.[16] Indeed, the way that the Local Authority was eventually criticised in court parallels Aquinas on humility. Judge Nicholas Marston and the senior social worker were both careful not to question 'the motivation of LA to do the right thing'.[17] Instead, their criticism focused on the 'misguided'[18] actions that were carried out towards that admirable purpose. In Aquinas's terms, the social workers aimed at a 'great thing'. Unfortunately, they did so 'against right reason'.

3.2. Humility about what can be known

Somerset demonstrates the need for humility about what is known, but epistemic humility has another aspect, humility about what it is possible to know in the circumstances. Another case, *A Local Authority v M ('M')*,[19] provides an example. The facts in this case are 'long and complex',[20] but a bare outline will suffice here. 'M' is a young man with a learning disability and autism. He lived at home until he was 18. After that, he was moved between his home and various residential placements.[21] The case concerned a series of allegations and counter-allegations between his parents and the Local Authority. The authority alleged that M's mother was excessively controlling, to an extent that caused M harm[22]; M's parents counter-alleged that the managers of the residential placements were both controlling and neglectful.[23] Although Mr. Justice Baker found overwhelmingly for the authority,[24] he made it clear that the case was evidentially complex,[25] and one evidential difficulty is relevant here. In the words of Dr. Carpenter, a consultant psychiatrist:

> 'At present, on any choice [M] makes, one is not totally confident that what is chosen is what he truly wants as his passive acceptance of most instructions makes it possible that, even if he did not really want what he chose, his continuation to use or eat or perform that

[5] [2014] EWCOP B25.
[6] *Ibid* [3].
[7] *Ibid* [6] (Judge Nicholas Marston).
[8] *Ibid* [8], [11].
[9] *Ibid* [6], [76].
[10] *Ibid* [10].
[11] *Ibid* [41].
[12] *Ibid* [45].
[13] *Ibid* [74].
[14] *Ibid* [83].
[15] *Ibid* [8], [10], [12].
[16] *Ibid* [40]–[41].
[17] *Ibid* [58].
[18] *Ibid.*
[19] [2014] EWCOP 33.
[20] *Ibid* [3] (Baker J).
[21] *Ibid* [1], [24], [31], [36], [39], [53].
[22] *Ibid* [2], [24], [60], [191].
[23] *Ibid* [31], [63], [68], [205].
[24] *Ibid* [211]–[219].
[25] *Ibid* [72]–[76].

choice could be evidence of his compliance as much as true desire'.[26]

In other words, attempts to determine M's non-coerced choices may have been experienced by M as further coercion. Here, a difficulty emerges for any analysis based narrowly on respect for autonomy and protection. The Local Authority and M's parents counter-alleged 'controlling behaviour'; so, as in *Somerset*, protection and respect for autonomy did not necessarily clash. Interferences with M's autonomy were what he was to be protected from. To protect M's autonomy, however, required some idea of when M was autonomous; and this could not be determined. Furthermore, what thwarted the investigation was the same thing as prompted it, the fear that a controlling environment led to M merely complying with authority figures.

Unlike *Somerset*, *M* does not just show a limit to what a potential intervener knew. It shows a limit to what they could know in the circumstances. This requires a different sort of response. A limit to what the intervener knows should usually prompt further investigations. In contrast, a limit to what can be known should prompt the abandonment of that line of investigation, the opening of other lines of enquiry, and great caution when acting. Even here, though, it is important not to let humility be confused with despairing inaction. Limits to knowledge in one respect and at one time need not be limits in every respect or forever. In *M*, Baker J endorsed the need for ongoing support to help M to articulate his own wishes.[27] In time, that might escape the problem that Dr. Carpenter identified. In the meantime, the judge also examined other evidence to determine which environments were likely to be unduly controlling. Both of these moves illustrate an important point. Even where epistemic humility leads to a recognition that something cannot be known, it need not lead to inaction. Instead, it can prompt *better* responses to suspected abuse. This is also true of another type of humility, humility about action.

4. Humility and action when responding to abuse

4.1. Humility about action

This section, like the last, examines humility when responding to a particular case of suspected abuse. This subsection addresses humility about action, respect for the limits to what an action can achieve. As with humility about knowledge, humility about action can guide a potential responder to abuse in ways that attention to autonomy and protection alone will not. The case of *The London Borough of Tower Hamlets v TB ('TB')*[28] illustrates this point. It concerned a 41 year old woman with a learning disability,[29] whose husband physically and verbally abused her.[30] Although the judgement also addresses TB's capacity to consent to sex and contact with her husband, and the question of depriving her of her liberty, it is the court's treatment of her residence that shows the need for humility about action. The reported case occurred about two years after TB had been placed in supported accommodation provided by the Local Authority, but it quotes at length an earlier interim judgment authorising that placement. This juxtaposition of two lines of reasoning separated by 2 years provides something very rare: a case in which we know both why a decision was reached in the Court of Protection and, in some detail, what happened next.

It is plain that a need to protect TB from 'the risk of physical violence' partially motivated moving her into supported accommodation.[31] It was not, however, the only factor in that decision. In the earlier hearing, TB was described as living an 'empty life' in 'a wholly oppressive environment where there is a sense of containment', which 'revolved only about …watching television'.[32] Furthermore, there 'were no skills being developed'.[33] These observations formed the basis of an argument from autonomy; emphasising the oppressiveness of TB's current life laid the foundations for an argument that moving her would enhance her ability to be self-determining. Indeed, Mr. Justice Mostyn was careful to disarm any countervailing argument based on TB's occasionally expressed desire to live in the marital home, saying 'the wishes and feelings of TB have varied and seem to be the product of the environment in which she is presently sited'.[34] On this reading, such wishes were not genuinely autonomous. In the interim judgment, then, it was found that a move into supported living would not only protect TB from her abusive husband but also enhance her autonomy. As in *Somerset* and *M*, autonomy and protection were made to harmonise.

In *Somerset*, protecting P and valuing her autonomy both seemed to require removing her from her home, but epistemic humility could have prompted a proper investigation of the cause of her bruising. Similarly in *TB*, both principles seemed to require removing TB from her home, but humility could have prompted a more sophisticated response. In this case, however, it was the limits of action, not knowledge, which did not receive attention. This is because a common asymmetry occurred in analysis. It was observed in some detail that TB's living situation did not afford her much autonomy. There was, however, no real analysis of whether the proposed action, moving her, was likely to increase her self-determination. Consequently, the limits to what could be achieved just by removing TB from a site of overt abuse were not taken into account. This led to professionals underestimating what a commitment to increasing TB's autonomy would entail. Two years after the initial decision, Mr. Justice Mostyn observed that she had 'not prospered there as much as everyone hoped. She seems to lead a rather isolated and lonely life, spending hours lying on the sofa watching TV'.[35] At that point, the need to provide her with additional support was belatedly recognised.[36]

TB does not provide as striking an example of moral error as *Somerset*, although 2 years of inadequate support should not be discounted. TB was, at least, protected from further domestic violence. It is, however, worth examining how she was protected. The focus on her lack of autonomy allowed the issue of protection to become distorted in a familiar way. As has been observed in other cases, it made the central problem not the abuser's actions but a purported 'defect' in the abused person's 'abilities to resist external pressure' (Pritchard-Jones, 2016: p. 60). The best way to respond to TB's husband's violence was therefore seen to be removing *her* from the house and depriving *her* of her liberty. Without rhetoric about promoting TB's autonomy this action looks, rightly, strange. This rhetoric, though, could be nothing more than hyperbole without humility about action. Only paying attention to the limits of any particular actions would have led to professional's recognising what TB needed to be self-determining. Once again, this criticism echoes Aquinas. To aim at increasing TB's autonomy was a 'great thing', but to think it could be so easily achieved was against 'right reason'. This example, too, deflates the stereotype of humility leading to inaction. Instead, it may have led to better action. Two years later, when it was recognised that TB needed more support, it probably did.

4.2. Humble actions

Humility about action is not the only connection between humility and action. There is also a need for humble actions. In other words, it is not enough for someone confronting suspected abuse to reflect inwardly on limits, whether of their knowledge or of their possible actions. They must also acknowledge those limitations to the people that

[26] *Ibid* [138].
[27] *Ibid* [138].
[28] [2014] EWCOP 53; [2015] 2 FCR 264.
[29] *Ibid* [1]–[2].
[30] *Ibid* [5(9)].
[31] *Ibid* [5], [23].
[32] *Ibid* [5(13)].
[33] *Ibid.*
[34] *Ibid* [5(14)].
[35] *Ibid* [8].
[36] *Ibid.*

they might affect. In the context of responses to suspected abuse, this will especially be the person thought to be at risk of abuse. The idea of humble actions is nothing new. Aquinas (1274/1947: IIa-IIae, 161, 3) recognises the value of 'outward acts' of humility performed with 'due moderation'. Three further points are, however, significant in the current context: humble actions can sometimes help to overcome the limits that they acknowledge; they raise the danger of false humility; and there is reason to suspect an endemic lack of this public humility in some services in England and Wales.

Humble actions can sometimes help to overcome the limits that they acknowledge. Although not directly concerned with abuse, two passages by judges in the Court of Protection exemplify this point. Each concerns the effect on the judge of going to hospital to meet a person that they might be making decisions for, and each has something of the character of a personal testimony. After visiting the person, Mr. Justice Peter Jackson says 'I obtained a deeper understanding of Mr B's personality and view of the world'[37]; and, similarly, Mr. Justice Mostyn 'the person I met was different in many respects to the person described in the papers'.[38] It is easy to misunderstand what has happened here. Because the person in question is being allowed to express a preference it might seem that the judge's actions can be completely understood by reference to autonomy. This is not so. Going to meet someone probably does help a judge to better respect that person's autonomy, but it does so because it is an expression of judge's awareness of their own limits. This expressive aspect of humility is important. When a High Court judge, with the real and symbolic power inherent in the role, leaves their courtroom to listen to someone with a mental disability, with all the stigma and history of that diagnosis, then an important reversal has taken place. Furthermore, by publicly acknowledging the limits to their knowledge of the person, these judges have allowed themselves to partially address that limitation.

Humility is necessary to address the limits that it acknowledges, but it is seldom enough. For meeting the person to make any real difference, the judges needed a bundle of further skills and virtues. For instance, they had to be approachable enough able to put the person at ease, attentive enough to hear their story, and open-minded enough to let their own assumptions be challenged. Still, humility is critical in two ways. First, without the initial act of expressive humility, the judge simply would not have had the opportunity to exercise the other skills. Second, many of these skills require some humility to develop. To be good at listening to others requires practice, but only those that have acknowledged they could be better at listening to others will actively engage with this practice. Although humility is not usually enough to address the limits it acknowledges, it can help create both the opportunity and the skills necessary to address those limits.

If humble actions can help to overcome the very limits that they acknowledge, then they can be surprisingly potent. This potency can, however, lead to dangers. If public displays of humility can be powerful, then there is a potential motive for false humility. The *Somerset* case provides a striking example:

> '… the family were invited to discuss plans about P's future and express their views. In fact it is clear that was not the reason they were invited at all. Far from a change of heart and an attempt to communicate the reason is clear. It was felt by Mr M on advice from the LA lawyers that: "The COP might pick up that no 'round table" meeting has been held and this might disadvantage us during the hearing"'[39]

Here, something that seemed to show that the Local Authority valued the families input, and therefore seemed to be based on some recognition that the 'experts' did not have all the answers, was a deliberate deception. The humility that the Local Authority was displaying was false in at least two ways. First, it was false in the direct sense that the professionals only seemed to acknowledge their limits. There was no genuine interest in learning from the family, just a desire to win a court case. Second, it was false in terms of audience. The act of holding a meeting appeared to address the family, but it was really addressed to the Court of Protection. False humility may be an endemic danger in social contexts where humility is valued. For example, in early modern England, where humility was highly prized, false humility seems to have been considered an almost ubiquitous threat (Clement, 2015: Ch. 2). Nevertheless, the danger of false humility should not be overstated. Judge Nicholas Marston was not deceived by the authority's false humility in *Somerset*, and it seems unlikely that the family were either. Indeed, far from false humility helping the professionals to get what they wanted, it eventually helped to condemn them.

There may be a lack of sufficiently humble actions in some services. This article is not the place for any systematic study of this phenomenon. All the same, it is easy to find a worrying example: professional behaviour around the premature deaths of people with learning disabilities. People with learning disabilities are more likely than others to die from causes that could be avoided with good-quality healthcare (Heslop et al., 2013: p. 90), and a lack of knowledge of the individual is often the critical factor in this process (Heslop et al., 2013: p. 83). Furthermore, the investigations needed for effective diagnosis are often neglected; and the concerns of people with learning disabilities, their families, and carers, are not taken seriously (Heslop et al., 2013: p. 57).[40] This suggests a lack of epistemic humility, a failure by professionals to acknowledge the limits of their knowledge. It also, however, suggests a failure to act humbly. Learning from someone with a learning disability, or from their family, often requires a professional to admit that in some areas these people have more expertise than they do. A failure to take their concerns seriously does not only reflect an inward failure to recognise limits. It is also a public failure to act with appropriate humility. Furthermore, this failure to act humbly even seems to persist after someone with learning disabilities has died prematurely. A review of one NHS Foundation Trust found that despite people with learning disabilities in contact with it having an extremely low average age of death, it systematically failed to investigate their unexpected deaths (Green et al., 2015: p. 16–17). Here, again, a failure of epistemic humility was intimately tied to a failure to act humbly; for instance, by listening to the concerns of families (Green et al., p. 32).

Reports into the deaths of people with learning disabilities raise extremely hard questions for professionals. People with learning disabilities consistently die young, and their early deaths have been consistently tied to professional failings (Heslop et al., 2013). Furthermore, the implicated patterns of professional behaviour continue even after someone with learning disabilities has died (Green et al., 2015). In these circumstances, it does not seem overly dramatic to ask whether, far from protecting people with learning disabilities from abuse, the state is systematically perpetrating abuse. If so, much of this abuse may be inflicted unwittingly. The professionals involved may, in a pattern familiar to social psychologists (Dunning, 2005: Ch. 2), not only fail to recognise their own limits but also fail to recognise that they have done so. This is another problem that the principles of autonomy or protection cannot solve alone. Before someone dies, a professional without humility can wrongly assume that they are that person's protector or that they empower the person to be autonomous; and, after someone dies, the same professional might conclude that there is now no-one that they need to protect or empower. Humility, in contrast, directs attention to the common source of the problem, the professional's unacknowledged limits.[41] In this way, once again, humility can help to

[37] *Wye Valley NHS Trust v B* [2015] EWCOP 60; [2015] COPLR 843 [18].
[38] *A Hospital NHS Trust v CD* [2015] EWCOP 74 [31].
[39] *Somerset* [70].

[40] Indeed, there is evidence that this failure to listen is an even more widespread problem (Beresford, Perring, Nettle, & Wallcraft, 2016: p. 16–18).
[41] A sufficiently arrogant professional might nevertheless think that they are excellent at humility. Hopefully, that thought is contradictory enough to worry all but the most narcissistic individuals.

shape interpersonal relationships: not only how someone directly responds to abuse but also potentially abusive institutional relationships. This ability to shape interpersonal relationships has implications for the more indirect relationships of the law.

5. Humility and the law

5.1. Humility and disability discrimination

The previous two sections show the need for humility when directly responding to abuse. This section explores the implications for law. First, this subsection shows that a lack of humility in responses to abuse will disproportionately affect people with mental disabilities, then later subsections examine the law's ability to foster humility. The discrimination point is important. The analysis so far only shows that a lack of humility when responding to abuse can be harmful, not that it will disproportionately affect those with disabilities. To see that it will, two complicating factors must be put to one side. First, the role of intention: in none of the cases discussed so far is there evidence that the professionals involved intended to discriminate against the person. Analysis in terms of humility will even concede this point. After all, saying that someone lacked humility admits that they aimed at a 'great thing'. A lack of discriminatory intent, however, does not mean that discrimination did not happen. 'Discrimination' in the UNCRPD includes discriminatory purpose *or* effect,[42] so unintentional discrimination remains discrimination.

The second factor complicating the link between humility and discrimination is the effect of domestic law, in this case the Mental Capacity Act 2005 ('MCA'). The examples in this article were decided under the MCA, but Section 2(1) of the Act limits its application to those who are unable to decide *because* of 'an impairment of, or a disturbance in the functioning of, the mind or brain'. It is probably impossible to construe this in a disability-neutral way, and this context may give the impression that any discrimination must be due to this law, and not due to a lack of humility. Such an impression would be misleading. Merely changing the law would not be enough to avoid substantive discrimination. If the relevant section of the MCA was removed, then the 'functional' test in section 3(1) of the Act, which requires that a person 'understand' and be able to 'use and weigh' the information relevant to a decision, would remain. Under this section, a higher proportion of individuals with mental disabilities than those without would be found to lack capacity (Flynn & Arstein-Kerslake, 2014). If that is so, though, then persons with disabilities would still disproportionately suffer the effects of insufficiently humble responses to suspected abuse.

Even if all perceived differences in understanding between people with and without mental disabilities are only an expression of stigma, then removing Section 2(1) of the Act would not be enough to avoid discrimination. Any judgement about someone's ability to understand, or use and weigh, information is necessarily influenced by the social context in which it is made (Banner, 2013); and widespread, strong, and unacknowledged negative attitudes towards people with disabilities seem to be part of that context (Wilson & Scior, 2014).[43] Stigma cannot simply be legislated away. A radical response to this problem might be to suggest the removal of all powers to intervene in cases of abuse. This, however, would achieve formal non-discrimination at the cost of substantive equality. People with mental disabilities suffer more abuse than those without (Hughes et al., 2012), so a wholesale refusal to intervene would disproportionately affect them (Kong, 2015). A more prudent response would be to suggest that that powers to intervene should be formally disability-neutral, effective, and carefully scrutinised. If, however, non-discrimination requires an effective general response to abuse, and an effective response to abuse requires humility, then legislative scrutiny must also pay attention to the need to foster humility.

5.2. Fostering humility with the law

At first, the idea of using legislation to foster humility can appear counter-intuitive. It is hard to imagine a law stating 'be humble', much less a judge ruling on whether or not a particular professional has complied with it. An important distinction must, however, be made: between calling for direct state enforcement of a virtue; and paying attention to the detail of how the law already promotes some virtues at the expense of others (Nedelsky, 2012: p. 69–73).[44] If humility is important, then it would be negligent not to do the latter. After all, the law already does influence professional responses to suspected abuse. If they sometimes seem insufficiently humble, then there is reason to ask whether the law sometimes fosters professional arrogance instead of humility.

The purpose of this article is to make the foundational point that humility is important when confronting suspected abuse, so it is not the place for a systematic analysis of the law of any jurisdiction. Nevertheless, the MCA provides a convenient illustration of the way that law can foster arrogance instead of humility, as the examples in earlier sections concerned its application. As is often the way with law, what the MCA omits is at least as significant as what it includes. Under the Act, a person's 'wishes and feelings' must be taken into account after it has been determined that they lack capacity to make a decision,[45] and the influence of the UNCRPD has caused this part of the Act to receive considerable attention (Martin et al., 2016: p. 39–42). The person's own perspective *during* a determination of incapacity has, however, received less recent attention. The Act does not allow incapacity to be assumed, inferred merely because of an unwise decision, or established merely by reference to someone's condition, behaviour, age, or appearance.[46] Inability to make a decision is, however, clearly taken to be an objective fact,[47] and the Act does not require any attention be paid to the person's own perspective on their capacity when determining this objective fact (Skowron, 2016).

If humility is important, then inattention to the person's perspective during capacity assessments is a problem. There are often significant limitations to how well someone can assess another person's understanding, especially on the short time-scale of many professional relationships. Indeed, sometimes the person being assessed might have a better grasp of their own understanding of a situation than the assessor does. Despite this, the law does not require that if someone disagrees with a capacity assessment, this be taken as evidence that the assessment might be wrong. Indeed, the opposite can happen, and disagreement can be taken as evidence that the person *lacks* capacity (Allen, 2009). In these situations, capacity assessments can be experienced by the assessed person as authority figures simply exerting power over them (Stefan, 1993: p. 782–5). When the assessed person may have been abused, then this is especially worrying, the repetition of a pattern that they may know all too well. Of course, there is no reason to think that this is what always, or even usually, occurs. Many capacity assessors are acutely aware of their limits, and correspondingly attentive to the person's perspective. That, though, is despite the law, not because of it.

[42] Art. 2.

[43] The science underlying the measurement of both bias and discriminatory outcomes has serious problems with both validity and reliability (Carlsson & Agerström, 2016). Nevertheless, absence of proof is not proof of absence, and the existence of widespread disability discrimination remains, unfortunately, very plausible.

[44] Nedelsky uses the language of 'values' not 'virtues'. The point is made here by analogy.

[45] MCA s4(6)(a).

[46] MCA ss1(2), 1(4), 2(3).

[47] MCA s2(1).

5.3. Legislative humility

The example of capacity assessments suggests that sometimes the law fails to foster humility, but it does not suggest that this process is inevitable. For example, the MCA could have required the person's own perspective be taken into account during assessment or included robust supports to contest a finding of incapacity. All the same, there are reasons to be cautious when suggesting legal reform. Passing legislation is an action; and, like any action, there are limits to what it can achieve. Legislative humility, respect for the limits of what law can achieve, is therefore needed. Indeed, some limits are intrinsic to the law. As Aristotle notes, 'all law is universal, and yet there are some things about which it is not possible to make correct universal pronouncements' (trans. 2002: 1137b13–15). In other words, no matter how certain and just a law may appear on paper, it will sometimes be uncertain or unjust when applied to a particular sets of facts. Sir Edmund Thomas may be right when he laments that this 'continues to elude so many in the legal community', whether practitioners or academics (2005: p. 116–117). If he is right, then there is a profound danger here. Those unaware of their limits characteristically also fail to realise that they are unaware of their limits (Dunning, 2005: Ch.2), so an unrealistic legal community may be largely unaware that it is being unrealistic. A lack of legislative humility risks the position, in Taylor's words, where 'we can't see any more the way these rules fit badly our world of enfleshed human beings, we fail to notice the dilemmas they have to sweep under the carpet' (2007: p. 742).

Taylor's warning is perceptive, but it is to be hoped that legal debate has not yet become so apocalyptically unrealistic. All the same, the need for legislative humility is real, as two quotations about the MCA, separated by almost a decade, illustrate. In 2005, Lord Falconer, moving for a second reading of the Bill that would become the Act, argued that it was necessary because 'the current law is confusing, often incomplete, and much misunderstood'.[48] The Bill duly became law. Nevertheless, after the Act had been in force for 7 years, a House of Lords Select Committee tasked with reviewing it concluded that it had 'suffered from a lack of awareness and a lack of understanding' (Select Committee, 2014: p. 6); and observed widespread confusion and incompleteness.[49] The MCA, it seems, did not even effectively clarify the law, much less achieve any more ambitious aim.

For the most part, the Select Committee blamed 'implementation' or 'culture', not the Act itself, for its problems. To acknowledge problems with implementation is, however, to implicitly acknowledge the limits of law. It directs attention to the other things, beyond the law, that the Act requires to have effect. Even so, the Select Committee underestimated the limits of legislation. It generally assumed, when talking about implementation, that the MCA had established a clear rule that the 'culture' should be brought into alignment with.[50] This is too optimistic. As Sunstein says, 'the content and nature of a legal provision cannot be read off the provision' (1995: p. 960). Interpreting a law requires social context. This, however, suggests that the idea that culture can be changed by bringing it into alignment with a rule is too simple, for the rule depends on its cultural context for its own meaning (McDowell, 1998). Law can affect culture, but it never does *this* by simply establishing a new rule; for the rule, if it is understood at all, is understood in the terms of the culture in which it is implemented. Instead of changing culture by setting standards, the law is, at most, one part of a slow and unpredictable conversation. This does not mean that legal responses to abuse cannot be improved. Legislative humility, like other forms of humility, does not prescribe inaction. It merely counsels scepticism about attempts to use the law to directly impose a new vision of society. This counsel is hardly new (for example: Oakeshott, 1947), but it does require a particular sort of ideal. That ideal is not the creation, then implementation, of a blueprint for the prevention of all abuse. Instead it is the organic growth of institutions characterised by 'attentive silence' (Weil, 1943/2005: p. 73), more concerned with understanding and responding to particular abuses than with grandiose initiatives from the institutional centre. In an era of international human rights law, this different ideal can seem perversely out of touch with the times.

6. Humility and the UNCRPD

6.1. The compatibility of the UNCRPD with humility

The previous section argues for legislative humility, awareness of the limits of the law; and notes an apparent conflict between such humility and international human rights instruments. In the case of the UNCRPD, this conflict could appear stark. After all, the Convention attempts to 'to promote, protect and ensure the full and equal enjoyment of all human rights and fundamental freedoms by all persons with disabilities'[51] across the world. This could be read as using the law to impose a new vision of society on the grandest possible scale. Furthermore, the scholarly reception of the Convention sometimes seems to reinforce this interpretation. It has almost universally hailed the Convention as a 'paradigm shift' in the relationship between disabled people and the law. This language captures something of the hope that attends the UNCRPD; but, where humility is concerned, it can be dangerous. It can gloss over important limits to understanding. Nowhere in the world is, on even the most generous interpretation, close to the sort of equality that the Convention promises. Accordingly, there is not yet a single example of how to achieve this equality or any concrete details of how it might work in any given society. In these circumstances, rhetoric about implementing a new paradigm too easily implies that the Convention has the answers. As with the Select Committee report into the MCA, the picture is of making a wayward culture align with a clearly established rule; and, once again, that is too optimistic. A more humble appraisal would be that, instead of having answers, the Convention poses the world a set of vitally important questions. This humble appraisal should not be mistaken for despair. Just because the questions posed by the Convention are not yet answered does not mean that they will not be in the future. Good answers to any question, however, are only usually found by those who actively search for them, and only those who admit that they do not yet know the details of how to make the promise of the Convention real are likely to join this search.

Worries about the language of 'paradigm shifts' are not enough to show that the UNCRPD is necessarily incompatible with legislative humility. In fact, the conflict between the two is more apparent than real, for the UN does not have the institutional machinery of a state. The Convention establishes a Committee on the Rights of Persons with Disabilities that can 'consider' allegations of breaches by states that are party to its Optional Protocol, but even this power is relatively weak. International human rights bodies depend on domestic systems to enforce their decisions, but States do not generally perceive these decisions to be binding, and most are not implemented (Alebeek & Nollkaemper, 2012). The Convention operates in a different context to domestic law, and this is significant. There is no danger of it being used to attempt to directly impose a new vision of society. It is primarily expressive, not directive (Lewis, 2010): a collection of 'standards and values' (Bartlett, 2014: p. 170) that can contribute to legal and social change in State Parties. Responsiveness is built into its institutional context. This is an example of a point made in Section 2, humility is especially needed by the powerful. UN human rights bodies do not have

[48] HL Deb 10 January 2005, vol 668, col 13.
[49] Paragraphs 63, 97, 167, 267, 276. Conclusions and Recommendations': paragraphs 31, 46, 48, 70.
[50] Paragraphs 15, 38, 39, 84, 85.
[51] Art. 1.

the power of states, so humility makes different, lesser, demands on them.[52]

6.2. Fostering humility with the UNCRPD

Beyond mere compatibility, the UNCRPD could help to foster humility when responding to suspected abuse. This might be a surprising claim. After all, the Convention does not mention 'humility', but 'autonomy' is one of its general principles,[53] and Article 16(1) requires 'measures to protect persons with disabilities'. The treaty seems to reinforce the established binary between autonomy and protection, not to challenge it. A careful reading, however, suggests that the Convention could help to foster humility about both knowledge and action. In both cases, this is because of the way that it treats autonomy. It does not treat autonomy as an objective feature of the person that professionals can promote by deploying their 'empowering' expertise. Instead, it makes a link between self-determination and being able to participate, and this undermines any easy assumption that professional expertise can be trusted.

The UNCRPD's potential to promote humility about knowledge becomes apparent when some passages on participation are examined. The preamble affirms that 'persons with disabilities should have the opportunity to be actively involved in decision-making processes about policies and programmes, including those directly concerning them'. Similarly, Article 3 makes 'full and effective participation and inclusion in society' a general principle of the entire Convention, and Article 12(4) requires that any safeguards relating to the exercise of legal capacity respect the 'rights, will and preferences of the person'. All of these Articles emphasise the perspective and expertise of the person with disabilities; and when the person's own perspective is in view, then the relative limits of expert knowledge are more obvious. When limits are more obvious, humility is likely to be easier. It is hard to imagine professionals taking these parts of the Convention seriously but nevertheless removing P in *Somerset* from her family, over her clear objections, without first fully investigating the cause of her bruising. If, then, the UNCRPD influences domestic systems and eventually individual practice, it could nurture epistemic humility. It will, however, only do so to a point. The emphasis on participation draws attention to the limits of what a professional can know, but all human knowledge is limited. There will also be limits to what the disabled person, or the disabled person and an expert working together, knows. Humility about knowledge demands more than the Convention.

The UNCRPD may also foster humility about action. Article 16, which requires appropriate measures be taken to prevent and respond to abuse, parallels the rest of the Convention by emphasising participation over expert intervention. In particular, Article 16(2) requires 'assistance and support for persons with disabilities and their families and caregivers, including through the provision of information and education'. The ideal here is not of a powerful expert intervening to 'save' a victim who has no agency, but of equals working together to avoid abusive situations from developing. By acknowledging that those at risk of abuse still have agency and power, the Convention, once again, makes the limits to expert power more obvious. Furthermore, this extends, beyond prevention, to responses to abuse. Article 16(4) states:

> 'States Parties shall take all appropriate measures to promote the physical, cognitive and psychological recovery, rehabilitation and social reintegration of persons with disabilities who become victims of any form of exploitation, violence or abuse, including through the provision of protection services. Such recovery and reintegration shall take place in an environment that fosters the health, welfare, self-respect, dignity and autonomy of the person'.

This part of the Convention could have made a difference to TB. To merely remove her from the house that she shared with her abusive husband, without providing further support, falls far short of these requirements. The Convention gives a substantive, albeit general, account of what a victim of abuse might require; and this makes it more difficult to ignore the limits of what a single intervention can achieve. Once again, however, humility demands more than the Convention. The Convention emphasises the limits of expert action; but there will also be limits to what the disabled person, or the disabled person and an expert working together, can achieve in any given circumstances.

The UNCRPD could help to foster humility about both knowledge and action when responding to abuse, but humility demands more than the Convention does. This is not surprising. The UNCRPD, although remarkably idealistic, is the product of political negotiation; but humility is something more rarefied, the product of ethical and theological reflection. There is a need for both of these things. Indeed, in the long term, one of the Convention's strengths may be that the hard questions that it poses prompt reflection even on the limits of human rights language. Sometimes, as when Onazi (2016) reads the Convention through Weil's concept of attention, this reflection can be deeply productive. It is also a reflection on *limits*; and so another, albeit unintentional, way that the Convention can foster humility.

7. Conclusion

Being humble is difficult. Medieval monks considered it difficult enough to require detailed guides, and it is no easier today. As Grenberg observes, 'to handle the fact of one's limit in a way that is admirable or exemplary is a slippery, difficult thing to do' (2005: p. 6). Anyone theorising about humility is well advised to remember one slippery difficulty in particular. The complexity of the human world means we must be content, as Aristotle says, with theories that are, at best, true 'roughly and in outline' (trans. 2002: 1094b13–28). In other words, social theories have limits, so no theorist of humility should expect too much from their work. The concept of humility will not end the abuse of disabled people or dispel all related ethical dilemmas. It will not even clear up all of the associated conceptual tangles. As Clement (2015: p. 25) says of 'Humilitie', George Herbert's allegorical poem, 'there is no decisive victory that would render Humilitie the hero'.

Although humility cannot win a decisive victory against abuse, there is reason to think that it might improve responses to abuse. This article shows how it could improve responses at an individual level; humility about knowledge and action can guide a responder when the principles of protection and respect for autonomy alone are inadequate. This has implications for legislation. If non-discrimination requires effective general responses to abuse, and effective responses to abuse require humility, then legislation should foster humility. Furthermore, the law has limits, and the legislator needs humility at least as much as the social worker. Finally, humility highlights the importance of participation as a general principle of the UNCRPD, and this underlines the differences between the Convention and more established approaches. Beyond that, humility can foster awareness of the limits of the treaty itself, and encourage treating it as a set of questions, not answers.

This article has many limits. It offers no proof that humility is important when responding to abuse, whether from first principles or by empirical study. Instead, it only offers a set of illustrative examples from one jurisdiction. It does not suggest how humility can be fostered when responding to abuse. Instead, it makes the preliminary point that, if possible, humility should be fostered. Finally, it does not make any detailed legislative proposals. Instead, it only indicates that legislative reform should both pay attention to humility and be humble. Some of these limits could be overcome by further work. Such work, however, would only be valuable if the central argument here, that humility is important when responding to abuse, is plausible. If it is not, then at least

[52] This does not show that UN human rights bodies *cannot* act in a way that is incompatible with humility, only that the Convention does not *require* it. For some worrying recent developments in human rights monitoring see MacGrogan, 2016.

[53] Art. 3.

the author, as Bernard (1120/1973: p. 28) says, 'will have nothing about which to be proud'.

Acknowledgments

I would like to thank Kirsty Keywood, Nicola Glover-Thomas, Laura Pritchard-Jones, and the anonymous reviewers for their constructive comments on earlier drafts of this article.

References

Alebeek, R. V., & Nollkaemper, A. (2012). The legal status of decisions by human rights treaty bodies in national law. In H. Keller, & G. Ulfstein (Eds.), *UN human rights treaty bodies: Law and legitimacy* (pp. 356–412). Cambridge: Cambridge University Press.
Allen, N. (2009). Is capacity "in sight"? *Journal of Mental Health Law*, *19*, 165–170.
Aquinas, St. T. (1947). *Summa Theologica. Translated by the fathers of the English Dominican Province*. London: Catholic Way Publishing (Original work published 1274).
Aristotle (2002). Nicomachean ethics. In Broadie, S. (ed) *Nicomachean ethics*. Translated by Rowe, C. Oxford: Oxford University Press. (Original work approximately 322 BC).
Augustine, St. (1997). Nature and grace. In Rotelle, J.E. (ed) The Works of Saint Augustine Volume 23 (pp. 197–268). Translated by Teske, R.J. New York, NY: New City Press. (Original work published 415).
Banner, N. F. (2013). Can procedural and substantive elements of decision-making be reconciled in assessments of mental capacity? *International Journal of Law in Context*, *9*, 71–86.
Bartlett, P. (2014). Implementing a paradigm shift: Implementing the convention on the rights of persons with disabilities in the context of mental disability law. *Torture in healthcare settings: Reflections on the Special Rapporteur on Torture's 2013 Thematic Report*. Washington, DC: Centre for Human Rights and Humanitarian Law.
Beresford, P., Perring, R., Nettle, M., & Wallcraft, J. (2016). *From mental illness to a social model of madness and distress*. London: Shaping our Lives.
Bernard of Clairvaux, St. (1973). *The steps of humility and pride*. translated by Conway, A. Spencer, MA: Cistercian Publications. (Original work published 1120).
Boyd, C. A. (2014). Pride and humility: Tempering the desire for excellence. In K. Thorpe, & C. A. Boyd (Eds.), *Virtues and their vices* (pp. 245–269). Oxford: Oxford University Press.
Button, M. (2005). A monkish kind of virtue?' for and against humility. *Political Theory*, *33*, 840–868.
Carlsson, R., & Agerström, J. (2016). A closer look at the discrimination outcomes in the IAT literature. *Scandinavian Journal of Psychology*, *57*, 278–287.
Clement, J. (2015). *Reading humility in early modern England*. Farnham: Ashgate.
Cooper, J. E. (2013). *Secular powers: Humility in modern political thought*. Chicago: University of Chicago Press.
Coulehan, J. (2011). A gentle and humane temper: Humility in medicine. *Perspectives in Biology and Medicine*, *54*, 206–216.
DasGupta, S. (2008). Narrative humility. *The Lancet*, *371*, 980–981.
Descartes, R. (1985). The passions of the soul. In *The Philosophical Writings of Descartes Volume 1* (pp. 325–404). Translated by Cottingham, J., Stoothoff, R., Murdoch, D. Cambridge: Cambridge University Press. (Original work published 1649).
Dunn, M. C., Clare, I. C. H., & Holland, A. J. (2008). To empower or to protect? Constructing the 'vulnerable adult' in English law and public policy. *Legal Studies*, *28*, 234–253.
Dunning, D. (2005). *Self-insight: Roadblocks and detours on the path to knowing thyself*. New York: Taylor and Francis.
Ehrlinger, J., Mitchum, A. L., & Dweck, C. S. (2016). Understanding overconfidence: Theories of intelligence, preferential attention, and distorted self-assessment. *Journal of Experimental Social Psychology*, *63*, 94–100.
Ellis, J. W. (1992). Decisions by and for people with mental retardation: Balancing considerations of autonomy and protection. *Villanova Law Review*, *37*, 1779–1809.
Flynn, E., & Arstein-Kerslake, A. (2014). The support model of legal capacity: Fact, fiction, or fantasy? *Berkeley Journal of International Law*, *32*, 124–143.
Green, B., Bruce, M., Finn, P., Wright, A., Daniel, D., Povey, J., & Repper, D. (2015). *Independent review of deaths of people with a learning disability or mental health problem in contact with Southern Health NHS Foundation Trust April 2011 to March 2015*. London: Mazars.
Grenberg, J. (2005). *Kant and the ethics of humility*. Cambridge: Cambridge University Press.
Heslop, P., Blair, P., Fleming, P., Hoghton, M., Marriot, A., & Russ, L. (2013). *The confidential inquiry into premature deaths of people with learning disabilities (CIPOLD)*. Bristol: Norah Fry Research Centre.
Hughes, K., Bellis, M. A., Jones, L., Wood, S., Bates, G., Eckley, L., ... Officer, A. (2012). Prevalence and risk of violence against adults with disabilities: A systematic review and meta-analysis of observational studies. *The Lancet*, *379*, 1621–1629.
Klancer, C. H. (2012). How opposites (should) attract: Humility as a virtue for the strong. *The Heythrop Journal*, *22*, 662–677.
Kong, C. (2015). The Convention for the rights of persons with disabilities and article 12: Prospective feminist lessons against the "will and preferences" paradigm. *Laws*, *4*, 709–728.
Lewis, O. (2010). The expressive, educational and proactive roles of human rights: An analysis of the United Nations Convention on the Rights of Persons with Disabilities. In B. McSherry, & P. Weller (Eds.), *Rethinking rights-based mental health Laws* (pp. 97–128). London: Hart.
MacGrogan, D. (2016). Human rights indicators and the sovereignty of technique. *The European Journal of International Law*, *27*, 385–408.
Martin, W., Michalowski, S., Stavert, J., Ward, A., Ruck Keene, A., Caughey, C., ... Mcgregor, R. (2016). Three jurisdictions report: Towards compliance with CRPD art. *12 in capacity/incapacity legislation across the UK. The Essex Autonomy Project*.
McDowell, J. (1998). Virtue and reason. *Mind, value, and reality* (pp. 50–73). Cambridge, MA: Harvard University Press.
Murdoch, I. (1970). *The sovereignty of good*. London: Routledge.
Nagel, T. (1986). *The view from nowhere*. Oxford: Oxford University Press.
Nedelsky, J. (2012). *Law's relations: A relational theory of self, autonomy, and law*. Oxford: Oxford University Press.
Oakeshott, M. (1947). Rationalism in politics. *Cambridge Journal*, *1* (81–95 &145–57).
Onazi, O. (2016). [Disability] justice dictated by the surfeit of love: Simone Weil in Nigeria. *Law and Critique*. http://dx.doi.org/10.1007/s10978-016-9191-2.
Parens, E. (1995). The pluralist constellation. *Cambridge Quarterly of Health Care Ethics*, *4*, 197–206.
Penrose, B. (2010). Humility and understanding. *Philosophical Papers*, *39*, 427–455.
Pritchard-Jones, L. (2016). The good, the bad, and the 'vulnerable older adult'. *Journal of Social Welfare and Family Law*, *58*, 51–72.
Richardson, G. (2007). Balancing autonomy and risk: A failure of nerve in England and Wales? *International Journal of Law and Psychiatry*, *30*, 71–80.
Select Committee on the Mental Capacity Act 2005 (2014). *Mental Capacity Act 2005: Post-Legislative Scrutiny. 2013–14, HL Paper 139*.
Skowron, P. (2016). The implications of Meno's paradox for the Mental Capacity Act 2005. *Medical Law Review*, *24*, 379–395.
Stefan, S. (1993). Silencing the different voice: competence, feminist theory and law. *University of Miami Law Review*, *47*, 763–815.
Sunstein, C. R. (1995). Problems with rules. *California Law Review*, *83*, 953–1026.
Taylor, C. (2007). *A secular age*. London: The Belknap Press of Harvard University Press.
Thomas, E. W. (2005). *The judicial process: Realism, pragmatism, practical reasoning and principles*. Cambridge: Cambridge University Press.
Weil, S. (2005). Human personality. In Miles, S. (ed) *Simone Weil: An Anthology* (pp. 69–98). Translated by Rees, R. London: Penguin. (Original work published 1943).
Williams, B. (2011). *Ethics and the limits of philosophy (Routledge Classics Edition)*. London: Routledge (Original work published 1985).
Wilson, M. C., & Scior, K. (2014). Attitudes towards individuals with disabilities as measured by the implicit association test: A literature review. *Research in Developmental Disabilities*, *35*, 294–321.