

**Public Reasoning and Reciprocity:
Reconceptualising Accountability in the Arm's Length
Governance of Emerging Biotechnologies in the UK**



Kyle Tierney Edwards

Thesis submitted for the degree of Doctor of Philosophy in Public Health

St Cross College

University of Oxford

Word Count: 83,890

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	iv
ABSTRACT	v
INTRODUCTION	1
CHAPTER I. THE LEGITIMACY PROBLEM: ACCOUNTABILITY AND INDEPENDENCE IN ARM'S LENGTH GOVERNANCE	12
1. <i>The current state of arm's length governance in the UK</i>	15
2. <i>The historical context of the quango debate</i>	24
3. <i>The legitimacy problem in the arm's length context</i>	44
CHAPTER II. KEY PLAYERS AND PRACTICAL CHALLENGES	48
1. <i>Human Fertilisation and Embryology Authority</i>	51
2. <i>National Institute for Health and Care Excellence</i>	64
3. <i>Medicines and Healthcare Products Regulatory Agency</i>	82
CHAPTER III. NAVIGATING THE PROBLEM: PUBLIC REASON AND DELIBERATIVE DEMOCRACY	96
1. <i>Deliberative versus aggregative conceptions of democracy</i>	100
2. <i>Public reason and the criterion of reciprocity: giving up on the 'whole truth' in politics</i>	106
3. <i>Procedure and substance in public reasoning</i>	114
4. <i>Deliberative democracy: the persistence of deliberative disagreement in spite of reciprocity</i>	118
5. <i>Can public reason yield definitive answers for arm's length bodies?</i>	124
6. <i>Where can we go when public reason 'runs out'?</i>	131
7. <i>A framework of public accountability as reciprocity for the arm's length regulation of emerging biotechnologies</i>	138
CHAPTER IV. RECIPROCAL REASONING IN THE HFEA: IMPLICATIONS FOR PUBLIC CONSULTATION, LAY REPRESENTATION, AND ETHICS EXPERTISE	146
1. <i>A third way: defending reciprocity against 'representation' and 'reason'</i>	149

<i>2. Implications of the reciprocal model for procedural requirements: public consultation and appointment of deliberators</i>	158
<i>3. The hard case of non-medical sex selection: what went wrong?</i>	174
CHAPTER V. THE ROLE OF PATIENTS IN NICE: PROVIDING AND CONTESTING EVIDENCE	189
<i>1. Voice or choice: delimiting patient involvement in technology appraisals</i>	197
<i>2. The patient voice as evidence</i>	202
<i>3. Contesting evidence: fostering the means of resistance in NICE appeals</i>	214
CHAPTER VI. REGULATORY BOUNDARIES IN THE MHRA: OPENING UP THE CONCEPTS OF SAFETY AND EFFECTIVENESS	232
<i>1. When is a product a 'medical' product?</i>	240
<i>2. Why is the medical sphere regulated uniquely?</i>	246
<i>3. Cognitive enhancement: benefit beyond 'medical' benefit</i>	253
<i>4. Telehealth systems: risk beyond 'health' risk</i>	260
<i>5. Implications for the processes and regulatory boundaries of the MHRA</i>	270
CONCLUSIONS	277
<i>1. Recommendations for Government: endorsing public accountability as reciprocity</i>	277
<i>2. Recommendations for arm's length bodies: applying public accountability as reciprocity</i>	285
<i>3. Contribution to the practical application of public reason and deliberative democracy</i>	291
<i>4. Implications for the concept of 'ethics expertise'</i>	295
<i>5. Implications for the Patient and Public Involvement movement</i>	297
<i>6. Future directions</i>	299
BIBLIOGRAPHY	301

Acknowledgements

First and foremost, I would like to thank my advisors, Michael Dunn, Julian Savulescu, and Imogen Goold, for their constant encouragement, incisive feedback, and good humour over the past three years, particularly in the final months of writing up this dissertation. I am also thankful for my wonderful colleagues at the Ethox Centre and the Uehiro Centre for Practical Ethics, with whom formal seminars and informal lunchtime conversations alike have greatly influenced the development of my thought on the problems considered here and the field of bioethics more generally. In particular, I am indebted to conversations with Mark Sheehan and Leah Rand that helped clarify the nature and goals of my research project in its earliest days.

In the larger Oxford community, I am exceedingly grateful to have found my way to the Foxes Graduate Football team during my first freshers' week. The friends that I have made there, especially my co-captains Rebecca Merkley, Alexandra Zeitz, and Lyndsey Hoh, have proved in turn silly, supportive, driven, and brilliant, and I cannot imagine the past three years without them.

My time at Oxford would not have been possible without the support of the Marshall Aid Commemoration Commission, and I have made lasting friendships and academic connections with my fellow Marshall Scholars in Oxford and across the UK. I feel extremely lucky to have been offered this opportunity 'across the pond,' and thank the Commission, especially its administrators Mary Denyer and Lizzie Clark, for making these amazing three years possible.

My parents, Chad and Ruth, and my brother, Drew, have offered constant and unconditional love and support during my time at Oxford, and I am deeply thankful to have such a wonderful family. My dad also found time to read and comment upon this dissertation (all of it!), and it is certainly clearer for his input.

Finally, Hugo Jackson has been perhaps the closest observer of my many ups and downs in completing this project, and has kept me smiling and laughing through it all. I could not have managed it without his constant good humour, love, and support.

Abstract

In the UK, public access to emerging biotechnologies is largely regulated by three government bodies that operate at arm's length from democratically elected representatives: the Human Fertilisation and Embryology Authority (HFEA), the National Institute for Health and Care Excellence (NICE), and the Medicines and Healthcare Products Regulatory Agency (MHRA). Purposefully situated beyond traditional lines of electoral accountability, how can these independent bodies (commonly referred to as 'quangos') conceptualize and demonstrate their accountability to the public in regulating technically complex and morally controversial technologies? After articulating this general theoretical tension between independent and accountable decision-making in the arm's length context, I illustrate how it generates distinct, practical challenges within the HFEA, NICE and the MHRA.

To navigate this general tension and address its practical ramifications, I invoke political theories of public reason and deliberative democracy, as well as recent attempts to apply these theories to the governance of biotechnologies. Through a critical analysis of these accounts, I articulate a new framework for 'public accountability as reciprocity' in the arm's length context. Its four, process-oriented conditions, of 1) reciprocal reasoning, 2) acceptance of deliberative disagreement, 3) voting, and 4) fostering the means of deliberative resistance, affirm the goal of identifying mutually acceptable reasons for collective decisions, but take seriously the fact that characteristics of emerging biotechnologies will often confound such attempts. Taken together, these conditions provide a more theoretically robust and normatively legitimate approach than existing accounts, notably 'accountability for reasonableness.'

Next, I translate the theoretical claims of this framework into practice, illustrating how its conditions may be fruitfully applied towards the resolution of the practical challenges identified within the HFEA, NICE, and MHRA. In the HFEA, public accountability as reciprocity provides a robust approach to moral reasoning in the body's deliberations and resituates the role of public consultation. In NICE, it suggests a justifiable and properly tailored role for patient involvement in technology appraisals. Finally, in the MHRA, it explains why a growing number of devices defy the institution's traditional assessments of safety and effectiveness and offers reforms to its regulatory boundaries.

I conclude with a set of policy recommendations for creating and operating arm's length bodies, as well as an account of broader contributions to the literature on public reasoning, the concept of 'ethics expertise,' and the practice of patient and public involvement.

Introduction

In 2006, less than thirty years after the first human birth using *in vitro* fertilisation (IVF), the UK public seemed largely to have accepted the use of novel assisted reproductive technologies to address infertility. Despite significant political and moral controversy about these new developments both at home and abroad throughout the 1980s, Parliament was able to establish the Human Fertilisation and Embryology Authority in 1990 to regulate clinical use of reproductive technologies, as well as embryo research.¹ Over time, and in part due to the HFEA's stewardship and attentive regulatory oversight, IVF came to be seen as routine medical practice, helping tens of thousands of citizens have children. Despite initial concerns and objections, then, IVF techniques became so widely accepted and commonplace that many "wonder[ed] at the initial alarm about them."²

However, on the cutting edge of scientific and medical innovation, new breakthroughs that promise substantial benefits but simultaneously raise serious public and ethical concerns always lie just beyond the horizon. The year 2006 saw the HFEA undertake a public consultation on *Choices and Boundaries*.³ The purpose of the consultation was to explore the range of public views about the use of pre-implantation genetic testing (PGD) for serious but lower penetrance, later onset, and potentially

¹ Duncan Wilson, "Creating the 'ethics Industry': Mary Warnock, in Vitro Fertilization and the History of Bioethics in Britain," *Biosocieties* 6, no. 2 (2011).

² Loane Skene, "Bioscience, Community Expectations and the Law," *The University of New South Wales Law Journal* 26, no. 3 (2003).

³ Human Fertilisation and Embryology Authority, "Choices and Boundaries Report: A Summary of Responses to the HFEA Public Discussion," 2006, http://www.hfea.gov.uk/docs/Choices_and_boundaries_Report_2006_summary.pdf.

treatable genetic conditions, like certain forms of breast and bowel cancer. Since 2001, the HFEA had permitted the use of PGD to select against serious and often fatal conditions present from birth, but its use for these lower penetrance, later onset conditions would substantially expand the number and range of conditions that clinics could offer to screen for future parents.⁴

The views expressed in this consultation illustrate the persistence and pervasiveness of fundamental moral disagreement within the UK about individual ‘choices’ and public ‘boundaries’ in the use of emerging reproductive technologies. On one side of the spectrum sit those still uniformly opposed to the use of PGD, whose views are often, but not always, religiously motivated: one respondent voiced the view that “no one has the right to decide if another life is worth living or not,” while another – identifying as a medical practitioner – asserted, “I reject the presupposition that any individuals with genetic conditions should be identified and selectively destroyed...regardless of risk.”⁵ Others, not necessarily opposed to the use of PGD for serious conditions, still worried about a slippery slope: “Specific learning difficulties like Asperger’s, dyspraxia, dyslexia, ADHD, etc, could be future candidates for PGD. This would be highly undesirable. It would look like eugenics or social engineering.”⁶ Finally, on the opposite end of the spectrum are those largely in favour of reproductive autonomy, like one respondent who argued, “What we [the public] think does not matter. It should actually be what the families think and at the end of the day they are the only people that can actually tell us how [the inherited condition] affects them.”⁷

⁴ Human Fertilisation and Embryology Authority, “Pre-Implantation - Information on ‘Saviour Siblings’ and Testing for Genetic Conditions,” last updated April 1, 2014, <http://www.hfea.gov.uk/preimplantation-tissue-typing.html>.

⁵ Human Fertilisation and Embryology Authority, “Choices and Boundaries Report: A Summary of Responses to the HFEA Public Discussion,” 8, 10.

⁶ Ibid., 13.

⁷ Ibid., 11.

As an internal HFEA report noted in presenting these consultation results, “For every question there was always a range of views representing very different opinions held by those who responded...it is clear that any Authority policy is unlikely to be welcomed by all groups.”⁸ This spectrum of perspectives speaks to the vibrant pluralism of UK society. It suggests widespread public disagreement not only about the moral permissibility of specific reproductive techniques, but also about what the relationship between government, society, and the individual ought to be in the response to rapidly developing, technically complex, and morally controversial biotechnologies. This underlying question has persisted even as the public has come to accept discrete developments, like IVF, suggesting its endurance and the importance of addressing it in relation to both existing and forthcoming innovations. In a liberal democratic society like the UK, how should the HFEA approach regulation in the face of this public, moral disagreement?

The HFEA is not alone in facing this regulatory challenge. Emerging biotechnologies also pose problems for the fair allocation of scarce healthcare resources by the National Institute for Health and Care Excellence (NICE), due to their expense and often minimal evidence base, and for the licensing of medical products by the Medicines and Healthcare Products Regulatory Agency (MHRA), due to their ‘transformative’ nature⁹ and failure to fit into existing regulatory classifications. Each of these regulators is a ‘quango,’ a governmental body set up at arm’s length from elected representatives to make decisions that are politically impartial and informed by the best expertise on the substantive issue at stake. This dissertation offers a novel approach to such arm’s length regulation of emerging biotechnologies – a framework of ‘public

⁸ Katy Berry, “Authority Policy on the Use of PGD for Lower Penetrance Inherited Cancer Susceptibility Conditions,” May 10, 2006, http://www.hfea.gov.uk/docs/PGD_-_choices_and_boundaries_May06.pdf.

⁹ Steve Clarke and Rebecca Roache, “Introducing Transformative Technologies into Democratic Societies,” *Philosophy & Technology* 25, no. 1 (2012).

accountability as reciprocity’ – that seeks to account for the existence of reasonable pluralism in the UK demonstrated by this brief, opening example.

Through analysis of both the historical development of arm’s length governance in the UK and its most recent large-scale political reform in 2010 by the Coalition government, Chapter I sets out the central ‘legitimacy problem’ facing quangos that regulate public access to emerging biotechnologies. Such arm’s length bodies are created for the express purpose of reaching *independent* and politically impartial decisions. As such, traditional conceptions of electoral accountability, and the democratic lines of accountability that elections cultivate, are inadequate to steer their decision-making. Nevertheless, we expect these bodies to be broadly *accountable* to the public: to seek out and act in the ‘public interest’ in creating binding regulations on behalf of society. This creates a practical tension within the arm’s length context between the goals of independent and accountable decision-making: how can we conceive of accountability in this context, given that quangos are purposely insulated from the most familiar method of being held to account in a democratic society? The challenge of navigating this tension is what I refer to as the legitimacy problem, the fundamental challenge that the dissertation seeks to address.

After articulating this general problem, I illustrate in Chapter II that it is more than an issue of mere theoretical interest: it manifests in practical challenges to decision-making currently facing arm’s length bodies in the UK. In this chapter, I describe the legal remit and institutional structure of the HFEA, NICE, and the MHRA, the three main bodies that regulate public access to emerging biotechnologies. With these basic institutional backgrounds in place, I show how the tension between political independence and public accountability generates confusion or uncertainty about how to approach decision-making within each of these three distinct contexts. The HFEA, in its

regulation of the use of assisted reproductive technologies and human embryo research, has been criticized, from the one side, for being too disconnected from the public perspective and, from the other side, for being too beholden to public opinion. It has struggled to locate a general and transparent approach to its decisions that represents a robust method of moral reasoning, whilst simultaneously incorporating and accounting for the results of its public consultations.

In contrast, NICE explicitly endorses ‘accountability for reasonableness’ as its approach to determining whether the NHS should fund new and expensive technologies, a set of ostensibly procedural conditions for legitimate decision-making that will receive a great deal of attention in later chapters. Despite implementing this developed framework for fairly allocating public healthcare resources, there is widespread confusion about what the role of patients within the technology appraisal process is, and should be. This confusion traces to the absence of a clear rationale for patient involvement in this context: given the need to keep in mind the broader public interest in the distribution of scarce resources, it is unclear what is required to be properly accountable to those most acutely affected by specific funding decisions.

Finally, the MHRA has long understood the demand of public accountability as one of ensuring that the medicines and medical devices that it licenses for use by citizens are both safe and effective. However, a growing number of new devices, such as cognitive enhancement devices, pose ‘borderline’ problems for the MHRA. These are cases in which there is practical uncertainty about whether a product should be regulated by the MHRA, and, if so, what it means for it to be ‘safe’ and ‘effective.’ Rather than representing mere outliers, these so-called borderline cases highlight the wider, institutional instability of some of the MHRA’s traditional regulatory boundaries and concepts, such as the distinction between the ‘medical’ and the ‘non-

medical’/‘consumer,’ challenging the conventional account of what it means to make a legitimate decision to license or restrict public access to drugs and devices.

Taken together, these first two chapters articulate the general problem I seek to address and the practical ramifications that motivate its consideration. To navigate this problem, I turn in Chapter III to the critical analysis of public reason and deliberative democracy, exploring whether and how concepts from these two strands of political theory can contribute to the normative development of my own framework for public accountability in the arm’s length regulation of emerging biotechnologies. Both sets of theories are based upon a deliberative (as opposed to aggregative) conception of democracy that denies the sufficiency of direct or representative voting in justifying politically coercive decisions. As they attempt to lay out non-majoritarian conditions for the legitimacy of collective rules, they suggest initial promise for the problem of making decisions that are simultaneously politically independent and publicly accountable.

Focusing on seminal accounts by John Rawls (on public reason), and Amy Gutmann and Dennis Thompson (on deliberative democracy), I identify ‘reciprocity’ as a defining and shared feature of the two approaches: the idea that coercive state decisions should be based upon ‘public’ or ‘mutually acceptable’ reasons that all those who must live by them can, at some level, accept. Although I endorse reciprocity, I go on to argue that, due to special characteristics shared by the class of emerging biotechnologies, reciprocal reasoning will not always – or even often – yield a determinate policy position. The nature of disagreement on the questions that emerging biotechnologies invoke – the start and end of personhood or the proper conception of distributive justice, for example – will often make the search for agreement on substantive ‘reasons,’ and how to weigh or order them, unrealistic. This suggests the importance of crafting procedural, rather than purely substantive, principles for publicly

accountable decision-making, as Norman Daniels and James Sabin attempt on similar grounds with their accountability for reasonableness approach.

However, in cases where deliberative disagreement persists to the time at which a decision must be made, such decisions will be based upon reasons that are not fully public or reciprocal in nature. With particular reference to accountability for reasonableness as well as the Nuffield Council on Bioethics' public ethics approach, I argue that past attempts to practically apply the demands of public reason and deliberative democracy have not taken this state of affairs seriously enough, either theoretically or normatively. Recognizing the persistence of reasonable disagreement requires a step beyond reciprocal reasoning or – in Daniels' and Sabin's terms – the search for 'relevant' reasons: that of actively fostering the means by which citizens may deliberatively contest decisions.

On the basis of this critique of both theories and practical applications of public reason and deliberative democracy, I propose at the close of Chapter III a process-oriented framework of 'public accountability as reciprocity' for the arm's length governance of emerging biotechnologies. Its four conditions of reciprocal reasoning, acceptance of deliberative disagreement, voting, and fostering the means of deliberative resistance have general implications for procedural requirements within arm's length bodies, such as the structure of deliberation, patient and public involvement, lay and expert membership, and public education, that I outline as well.

The next three chapters translate the theoretical claims of this general framework to the specific contexts of the HFEA, NICE, and MHRA in order to resolve the practical challenges set out in Chapter II. Crucially, the framework does not provide a simple, one-size-fits-all solution to the legitimacy problem. Understanding the procedural requirements that follow from its application demands careful attention to institutional

context and the nature of the regulatory decisions in question. Therefore, while the resolution of the challenges identified in Chapter II is a valuable contribution of the dissertation in and of itself, the three chapters serve more broadly to illustrate how public accountability as reciprocity can be applied fruitfully in three, quite diverse settings. As such, I refer to Chapters IV, V, and VI as ‘case study’ chapters, in that they allow me to illustrate the practical usefulness of my framework and suggest its versatility in different kinds of contexts and for different types of biotechnologies.

In both Chapter II and these three case study chapters, I draw on a wide range of publicly-accessible documents from Government, Parliament, the European Commission and the HFEA, NICE, and the MHRA, including: primary legislation, statutory instruments, European Commission directives, background institutional documents, meeting minutes, annual reports and strategic plans, published internal working papers, consultation documents, policy decisions and technology appraisal determinations, Department of Health and Cabinet Office guidance and reports, independent external reviews, parliamentary evidence and debate, and parliamentary Select Committee reports. Alongside analysis of published empirical research studies, this documentary evidence permits the construction of full and rich accounts of both institutional context and the specific policy decisions that I refer to in the case study chapters. Careful and comprehensive review of this wealth of evidence meant that it was not necessary to conduct primary empirical research of my own in order to construct these reliable accounts. This is fortuitous, given the practical difficulties of accessing members of arm’s length bodies: appointed members almost always have other full-time careers, and thus have limited time to participate in research, and the size of decision-making committees is usually limited to allow for real deliberation amongst members, meaning the sample size from which to draw is small. Furthermore, spending cuts across

Government since 2008 mean that arm's length bodies are often asked to carry out more public functions with fewer staff and less funding, suggesting limited institutional resources for engaging in external research projects.

In Chapter IV, to address the HFEA's lack of a robust 'method of reasoning,' I propose a 'reciprocal method' and defend it against arguments that a public body making value judgements like the HFEA should either 1) seek to reflect public views and values or 2) seek to identify the position most capable of withstanding philosophical debate. I illustrate the implications of this reciprocal method of reasoning for the practice of public consultation and the appointment of lay members and philosophers to the HFEA. Finally, I use this method as a lens with which to critique the HFEA's process of decision-making in one of its former cases, that of non-medical sex selection, suggesting where it fell afoul of public accountability as reciprocity and outlining recommendations for improvement.

Guided by Daniels' and Sabin's accountability for reasonableness approach, the NICE Board already seeks to identify reciprocal reasons that should form the basis of its 'social value judgements,' and uses public input from its Citizens Council in this process. Technology Appraisal Committees are in turn bound by these values, which include clinical effectiveness, cost effectiveness, and non-discrimination, in their decision-making. As such, these Committees face a very different manifestation of the legitimacy problem than that examined in the HFEA: they must determine what the *substantiation* of these general, pre-determined values requires in the consideration of specific treatments and products.

Given that these substantive *values* have been previously (and properly) set in a publicly accountable manner, I argue in Chapter V that the only role that patients can justifiably play as consultees within this process is as providers of *evidence*. I illustrate

how patients can offer unique information that addresses limitations of clinical trial and economic data, such as the range of factors affecting their quality of life, and suggest how NICE can reform its patient involvement processes to better align with this rationale. However, recognising patient involvement as evidence provision, while resolving the ambiguous role for patients in technology appraisals, creates a second practical challenge for NICE: how to weigh this type of evidence against traditional clinical and cost effectiveness data, and how to handle reasonable disagreement about its interpretation. Through a review of NICE's appeals decisions, focusing on appeals that alleged that the Appraisal Committee's decision was 'unreasonable' in light of the evidence considered, I suggest how reasonable disagreement about evidence can and should be handled, by reference to my framework's condition of fostering the means of resistance.

In Chapter VI, the final case study chapter, I return to the MHRA's borderline problems in regulating new developments in the devices field. The MHRA has long understood and publicly presented its regulatory decision-making as purely '*fact-based*,' and has accordingly attempted to address these borderline problems by bolstering its access to clinical expertise in the area of medical devices. However, I argue that borderline devices defy the MHRA's traditional regulatory approach not because of any lack of technical expertise but because they throw into sharp relief the significant *value* judgements implicated by its measurement of risks and benefits and its determination of 'acceptable' trade-offs between the two.

To solve borderline problems, then, the MHRA must recognize and open up these value judgements to public deliberation and reciprocal reasoning, in which these evaluations are informed by the perspectives of the patients or users on whose behalf they are made. By considering cognitive enhancement devices and telehealth devices

like GPS trackers and fall detectors, I illustrate that the range of considerations relevant to the MHRA's risk-benefit calculations is far broader than the health risks and benefits to which it currently limits itself. Furthermore, individuals may value these risks and benefits in different, but nonetheless reasonable, ways. On the basis of these findings, I suggest how the MHRA can adjust its regulatory boundaries and processes to incorporate such considerations in a publicly accountable manner.

I conclude by enumerating formal recommendations for the Government, the HFEA, NICE, and the MHRA, as well as suggesting the implications of my work for other arm's length bodies involved in the regulation of emerging biotechnologies, both within the UK and abroad. Beyond these policy recommendations, I also illustrate the contributions of public accountability as reciprocity to the literature on the practical application of public reason and deliberative democracy, the concept of ethics expertise, and the practice of patient and public involvement.

Chapter I. The Legitimacy Problem: Accountability and Independence in Arm's Length Governance

Like the governments of two of his predecessors, Margaret Thatcher and Tony Blair, David Cameron's Coalition government was partially ushered in on a platform of quango reform. Capturing a common criticism of these 'quasi-autonomous non-governmental organizations' little less than a year before taking office as Prime Minister in May 2010, Cameron noted in a speech to the think tank Reform, "Too many state actions, services and decisions are carried out by people who cannot be voted out by the public, by organizations that feel no pressure to answer for what happens – in a way that is completely unaccountable."¹⁰ Indeed, unaccountability is just one in a long and interrelated list of offences of which quangos are frequently accused, including inefficiency, ineffectiveness, wastefulness, political patronage, and opaqueness.

The defining feature of quangos as an institutional class is their 'arm's length governance,' so-called because they are granted varying degrees of independence from their 'sponsor' governmental department, such as the Department of Health, and "accordingly operate at a greater or lesser extent at arm's length from ministers."¹¹ Many of the objections to quangos raised by politicians, the media, academics, and the public are intricately tied to this distinctive feature, particularly concerns about accountability, transparency, and independence. In one of the most thoughtful and theoretically rigorous

¹⁰ "Tories Pledge to Cut Back Quangos," *BBC News*, July 6, 2009, http://news.bbc.co.uk/go/pr/fr/-/1/hi/uk_politics/8135460.stm.

¹¹ Cabinet Office, "Public Bodies Reform," last updated November 10, 2014, <https://www.gov.uk/public-bodies-reform>.

treatments of arm's length governance in the UK, Matthew Flinders suggests that the British *ad hoc*, cavalier approach to state-building and delegated governance, which he terms 'walking without order,' provokes challenges in these areas that are far more pronounced than in other countries that practice delegated governance (synonymous for our purposes with arm's length governance).¹² Crucially motivating the analytic approach of this dissertation, he notes in the final chapter of *Delegated Governance and the British State: Walking Without Order*:

[S]cholars have generally focused their attention on the secondary consequences of delegation (the *politics* of delegation) instead of examining the roots or origins of the logic of delegation, or its permeation across governing regimes. Indeed, the dominant rationality in favour of widespread delegation portrays this process as a neutral element of 'good governance' when in fact it clearly masks the redistribution of political power within and beyond visible democratic frameworks. The logic of delegation itself, rather than its consequences, needs to be the focus of critical political analysis in order to push the process back within the sphere of public contestation.¹³

He observes, in other words, that the study of arm's length governance by quangos has tended to focus on its effects rather than on its rationale and, more critically, whether and how this rationale is publicly justifiable once granted proper consideration.

This trend has obscured a pair of questions relating to the normative underpinnings of arm's length governance: 1) which decisions should be taken at arm's length from democratically elected officials, and 2) how those decisions placed at arm's length should be made. By illuminating these oft concealed questions, this chapter develops an account of the central 'legitimacy problem' posed by delegated governance: given the political independence that follows from setting a decision at arm's length, the standard conception of democratic accountability is ill-suited for either understanding or

¹² Matthew Flinders, *Delegated Governance and the British State: Walking without Order* (Oxford: Oxford University Press, 2008): 311.

¹³ *Ibid.*, 324.

maintaining the normative legitimacy of arm's length bodies. This theoretical tension between independence and accountability results in practical confusion about what it means, or what is required, for an arm's length body to make a legitimate public decision. Each of the three institutions that regulate public access to emerging biotechnologies – the HFEA, NICE, and MHRA – falls within this broad class of arm's length bodies. And, as I will argue in Chapter II, this general legitimacy problem generated by placing public functions at arm's length manifests itself in specific decision-making challenges facing each body, such as uncertainty about the role of patient and public involvement or the justifiable exercise of expertise. As such, any attempt to address these challenges requires attention to the distinctive characteristics of this institutional form, the arm's length body, and the political issues it raises.

Responding to Flinders' call for a focus on the logic or rationale of delegation, this chapter aims to tease out the central considerations that arm's length governance poses for regulatory decision-making in the UK. Section 1 analyzes the most recent attempt, by the Coalition government, to reform the quango landscape, concluding that this reform programme failed to provide a conception of accountability *within* the arm's length context, resulting in persistent, practical confusion about how public decisions should be made by such bodies. Section 2 situates this current state of affairs within the broader history of arm's length governance in the UK, in order to provide a fuller account of the problems that thus arise. I pay particular attention to 1) rhetorical shifts in the political objections to quangos, 2) successive attempts in the literature to solve the terminological confusion surrounding arm's length governance, and 3) the evolution of the accounts of – and relationships between – accountability, independence, transparency, and legitimacy in the arm's length context. Finally, Section 3 provides a decontextualized account of the central problem that decision-making at arm's length

poses for regulating in the ‘public interest’ in a modern democracy like the UK. This sets the stage for the following chapter’s introduction of the three arm’s length bodies that regulate public access to emerging biotechnologies and how this general legitimacy problem manifests in each of these practical, decision-making contexts.

1. The current state of arm’s length governance in the UK

This section must begin with a brief note on terminology, although the topic will receive more attention in Section 2. While the acronym ‘quango’ is the most widely known term for public bodies that operate at arm’s length, it is perhaps the least useful in a set of unwieldy terms. As Tom Gash and Jill Rutter pointedly explain:

‘[Q]uasi-autonomous’ is hardly conducive to public understanding and ‘non-government’ is nonsense as these are bodies performing government functions, appointed by ministers, sponsored by departments and usually funded by them. The official titles are not much more helpful. Non-departmental public bodies report to departments; non-ministerial departments often both have ministers and are in the words of one chief executive ‘not real departments.’ Even the names of the bodies themselves appear almost designed to confuse us. The Environment Agency is an ‘executive NDPB’, the Food Standards Agency a ‘non-ministerial department’ and neither are ‘executive agencies.’¹⁴

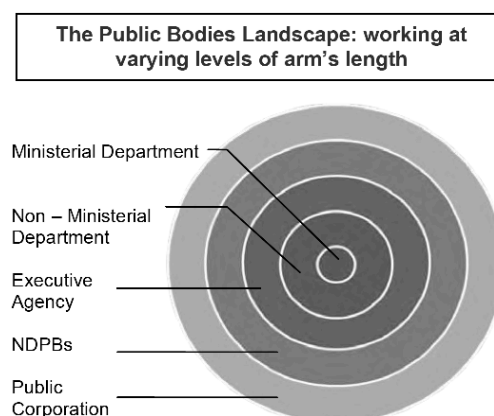


Figure 1. Cabinet Office, *Public Bodies 2012* (2012).

¹⁴ Tom Gash and Jill Rutter, “The Quango Conundrum,” *The Political Quarterly* 82, no. 1 (2011): 96.

Because the defining feature of these public bodies is their institutional location at arm's length from elected officials, I refer to them on the most general level as arm's length bodies. This is in keeping with much of the academic literature on the subject, although *quango* continues to be the preferred, informal term of politicians and the media.

There are four types of arm's length bodies: non-ministerial departments, executive agencies, non-departmental public bodies (NDPBs), and public corporations (see Figure 1). Although the MHRA is an executive agency and the HFEA and NICE are both NDPBs, a brief explanation of all four types is useful for contextualizing these two key ones. Non-ministerial departments are formally the 'closest' to democratic control, as they are accountable to Parliament through their sponsoring minister, and – as departments – they are funded directly by Parliament. Well-known examples include HM Revenue & Customs and the Food Standards Agency. Furthest from ministerial control are public corporations, which are bodies operating commercially that recoup at least 50 percent of their costs through commercial revenue. The British Broadcasting Corporation and the Disclosure and Barring Service are both public corporations.

Operating in the grey zone between these two categories are executive agencies and NDPBs. Executive agencies are structurally distinct from NDPBs because they are formally part of a ministerial department: each is funded by a parent department whose minister is directly accountable to Parliament for the executive agency, although most day-to-day operations are independent of the department and minister. Importantly, executive agencies are intended to be “well-defined business unit[s] with a clear focus on delivering specified outputs within a framework of accountability to ministers.”¹⁵ So while they have more autonomy than persons or groups working directly within a

¹⁵ Cabinet Office, “Public Bodies Reform.”

department, their remit is still to determine *how* to accomplish a given policy goal rather than *what* that goal ought to be in the first place.¹⁶

In contrast, NDPBs are not formally part of any department, although each has a sponsor department and is ultimately accountable to it and the sponsor department's minister. NDPBs are numerous (at last count, 450¹⁷) and incredibly diverse. There are four types of NDPBs: executive NDPBs, advisory NDPBs, tribunal NDPBs, and independent monitoring boards of prisons and immigration centres. The latter two can be left aside for our purposes as they occur rarely and in specific circumstances that are far from the field of biotechnology. Executive NDPBs are generally established by statute and "carry out a wide range of administrative, commercial, executive and regulatory or technical functions which are considered to be better delivered at arm's length from ministers."¹⁸ Advisory NDPBs, on the other hand, are "set up to provide independent, expert advice to ministers on an ongoing basis."¹⁹ The key difference to note is the way in which each type influences policy: even if an advisory NDPB makes a policy recommendation to a minister or other government official, it requires the approval of a separate, executive decision-maker to become policy, whereas many executive NDPBs both create and enforce policy independently. Both executive NDPBs and advisory NDPBs – albeit via different routes – contribute to policy making around science and medicine. However, the two NDPBs that I consider as gatekeepers to novel biotechnologies, the HFEA and NICE, are both executive NDPBs.

¹⁶ The discussion of standard-setting and standard-using in Chapter III will return to the importance of this distinction.

¹⁷ Cabinet Office, "Public Bodies 2014: Summary Data," June 3, 2015, <https://www.gov.uk/government/publications/public-bodies-2014/public-bodies-2014-summary-data>.

¹⁸ Cabinet Office, "Categories of Public Bodies: A Guide for Departments," December 2012, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/80075/Categories_of_public_bodies_Dec12.pdf, 15.

¹⁹ Ibid., 18.

Types of Arm's Length Bodies	Sub-classifications
Non-ministerial departments	
Executive agencies	
Non-departmental public bodies (NDPBs)	Executive NDPBs, Advisory NDPBs, Tribunal NDPBs, Independent monitoring boards
Public corporations	

Table 1. Taxonomy of Arm's Length Bodies.

Note that even some of the most basic claims in this initial discussion of terminology are contentious. Many would include more, or fewer, types of bodies in the taxonomy presented in Table 1. Likewise, some would challenge the practical reality of the theoretical distinctions suggested here between the various bodies. A proper understanding of these disputes is essential to comprehending the issues posed by arm's length governance and accordingly will be analyzed in Section 2. The definitions and distinctions set out here should thus be understood as necessary but rough guides.

A. The Coalition government's public bodies reform programme

Both Margaret Thatcher and Tony Blair, in 1979 and 1997 respectively, called for a 'bonfire of the quangos' leading up to and in the aftermath of their elections.²⁰ Proving no exception, David Cameron's Coalition government began a reform of the more than 900 'public bodies' in existence shortly after the 2010 election. This included 679 NDPBs and 222 other non-ministerial departments and public corporations. In October 2010, Francis Maude, the Minister for the Cabinet Office, announced the initial proposals for reform following this massive review: more than half the bodies were to be substantially reformed, nearly 200 abolished, and over 100 merged. Furthermore, all

²⁰ Gash and Rutter, "The Quango Conundrum."

remaining public bodies would be subject to similar ‘triennial reviews’ in the future.²¹ In defending both his method of review and proposed reforms, Maude claimed:

This public expect Ministers to take responsibility for what the Government do, and not leave this to people or bodies that are unelected. In the past, too many public bodies have been established without proper thought, and allowed to remain when their mission has long been accomplished. This has meant that elected politicians have been able to avoid making difficult and tough decisions. This is a direct challenge to accountability and is contrary to openness and transparency in public services that this Government seeks to achieve.²²

Because of this perceived accountability deficit, the Coalition government’s guiding presumption in its reform of quangos was that “if a public function is needed then it should be undertaken by a body that is democratically accountable at either national or local level.”²³

Based on this presumption of democratic accountability, Maude’s criteria for the reform of arm’s length bodies consist of a two-step method, each of which is focused on the nature of the function that the body performs. The first step is to determine whether a given function “should continue to be delivered at all,” sometimes referred to as the ‘existential test.’²⁴ If this first step is met affirmatively, the second step is to determine whether that function should remain at arm’s length from democratically elected ministers. This second stage requires that all arm’s length bodies meet at least one of three tests for whether arm’s length governance is appropriate: 1) it performs a technical function, 2) its activities require political impartiality, or 3) it needs to act independently to establish facts.²⁵

²¹ HC Deb 14 October 2010 vol 516 col 26WS – col 28WS.

²² Ibid.

²³ Cabinet Office, “Public Bodies Reform.”

²⁴ Cabinet Office, *Government Response to the Public Administration Select Committee Report, “Smaller Government: Shrinking the Quango State,”* (Cm 8044, March 2011), 9.

²⁵ HC Deb, 14 October 2010, vol 516, col 26WS – 28WS.

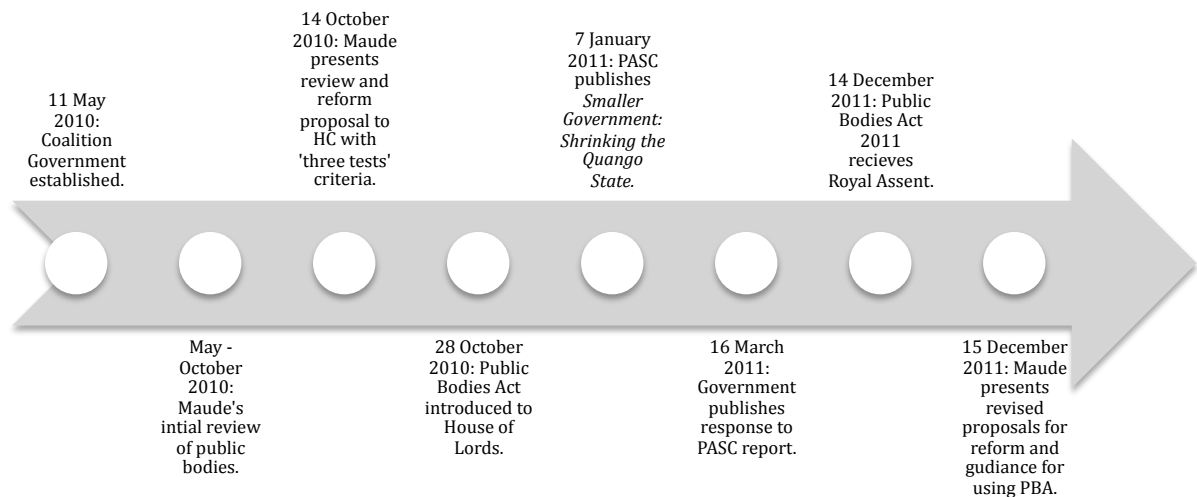


Figure 2. Timeline of Coalition Government's review of public bodies.

As the Public Administration Select Committee (PASC) argued in its report on the government's plans, these second-stage tests "may have seemed superficially plausible at the outset, but they are hopelessly unclear."²⁶ PASC claimed that the "common denominator for all three tests is independence from government," and thus suggested that the three tests could and ought to be collapsed into one test for whether the function requires impartiality.²⁷ Nevertheless, the Coalition Government rejected this recommendation in its response to PASC,²⁸ ultimately reifying the 'Maude tests,' and the requirement of triennial reviews using this methodology, in its guidance to departments for using the *Public Bodies Act 2011*.²⁹ The *Public Bodies Act*, which received Royal Assent on 14 December 2011, is legislation that allows ministers to abolish, merge and

²⁶ Public Administration Select Committee, *Smaller Government: Shrinking the Quango State*, 7 January 2011, HC 537-I 2010-2011, para 19.

²⁷ *Ibid.*, para 18.

²⁸ Cabinet Office, *Government response to the Public Administration Select Committee Report, "Smaller Government: Shrinking the Quango State."*

²⁹ Cabinet Office, "Using the Public Bodies Act 2011: A Guidance and Best Practice Note for Officials," September 2013, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/243888/Using_the_Public_Bodies_Act_2011.pdf.

reform public bodies that were created by an act of Parliament and which thus would typically require parliamentary approval to alter.

B. A gap in the reform programme

For arm's length bodies that do satisfy the admittedly vague Maude tests – and thus are neither abolished nor merged upon review – there exists little helpful guidance for internal review and reform. The Cabinet Office advises:

...the Department, with the Chair and CEO of the NDPB concerned, should then review the control and governance arrangements in place to ensure that the public body is operating in line with recognized principles of good corporate governance. This includes requirements on openness, transparency and accountability.³⁰

These 'principles of good corporate governance' as outlined by the Cabinet Office focus largely on statutory requirements, conflicts of interest in board and executive appointments, and financial management.³¹ Where the principles do provide advice on the process of decision-making within NDPBs, it is not especially action-guiding. For instance, the 'Communications Principle,' requires that "the Public Body is open, transparent, accountable and responsive," which is in turn supported by provisions to "identify its key stakeholders" and "engage and consult with the public on issues of real public interest or concern."³² Principles and provisions for the 'Role of the Board' suggest that the board and its committees "have an appropriate balance of skills, experience, independence and knowledge" and that the Board "establish appropriate arrangements to ensure that it has access to all such relevant information, advice and

³⁰ Cabinet Office, "Guidance on Reviews of Non Departmental Public Bodies," June 2011, http://www.civilservice.gov.uk/wp-content/uploads/2011/09/triennial-reviews-guidance-2011_tcm6-38900.pdf, 7-8.

³¹ Ibid., Annex B.

³² Ibid., Annex B, 32.

resources as is necessary to enable it to carry out its role effectively.”³³ While such abstraction may be necessary for addressing such a diverse set of institutions, the current guidance avoids questions fundamental to the operation of arm’s length bodies, particularly those that regulate emerging biotechnologies, like who constitutes a ‘stakeholder,’ what it means to ‘engage and consult,’ what information, advice, and knowledge is ‘relevant,’ and, perhaps most crucially, what issues are of ‘real’ public interest.

As a result, arm’s length bodies that survived the most recent bonfire – a political act ostensibly based on the underlying accountability problems of delegated decision-making – are ironically left with very little guidance on what decision-making at arm’s length should look like, or how it might be reformed to address these accountability concerns. These ‘survivors’ include the HFEA and NICE, both executive NDPBs. Furthermore, many advisory NDPBs in this area have simply been folded into executive agencies as ‘independent advisory committees’ that largely retain their structure and function. For instance, many such committees now operate within the framework of the MHRA, which is an executive agency of the Department of Health. This includes the Commission on Human Medicines, the Committee on the Safety of Devices (soon to be replaced by the Devices Expert Advisory Committee), and the Advisory Board on the Registration of Homeopathic Products, all of which were previously distinct NDPBs that the Coalition government classified as abolished. Although they now formally sit a bit ‘closer’ to ministerial control as they have been ‘hived in’ to an executive agency, it is far from apparent that the way in which they operate to produce decisions and influence policy has changed. Since executive agencies were classed out of Maude’s review, the incorporation of these former NDPBs means that their functions will continue to be

³³ Ibid., Annex B, 24-25.

overlooked in the context of arm's length body reform, as will that of the MHRA more generally.

This lack of guidance for decision-making at arm's length is surprising. Even if we agree that some public decisions ought to be made without undue influence from current politicians or on the basis of certain kinds of expertise, this need not be equated with the conclusion that there ought to be *no* guiding principles for how such public decisions should be made in a democratic society. Put otherwise, even if we place decisions at arm's length, and thus beyond traditional lines of democratic accountability, we may nevertheless still expect such bodies to behave in a manner that is publicly accountable and in line with the values underpinning democratic governance. Yet with regards to those arm's length bodies that pass the three Maude tests and exhibit 'good corporate governance,' neither the current Conservative government nor the preceding Coalition government has provided any further requirements, let alone guidance, for their functioning.

A number of think tanks have criticized the Maude tests and offered alternative ways of thinking about arm's length governance reform. For instance, the Institute for Government, in a proposal lauded by the Public Administration Select Committee, has suggested that Maude's three tests should be folded into one test based on "the degree of independence from day-to-day ministerial intervention needed to enable the body to *command public confidence that it can perform its function in the public interest*" (emphasis added).³⁴ This expressly links the design of arm's length governance back to the public interest, a notable improvement on the Coalition government's formulation.

³⁴ Tom Gash et al., "Read before Burning: Arm's Length Government for a New Administration" (Institute for Government, July 2010).

Yet even this test would require a better practical understanding and elaboration of what it means to perform a function ‘in the public interest.’

By limiting itself to reform based on either abolition or the movement of functions ‘closer’ to democratic control, the Coalition government failed to engage with the concept of accountability *within* the arm’s length context – somewhat ironic due to the programme’s ostensible aim of addressing accountability problems. Silence on this question entails a substantial delegation of political authority not only on the content of specific decisions but also on the question of *how* public decisions ought to be made, whenever those decisions require substantial 1) technical expertise, 2) political impartiality, or 3) independence in establishing facts. As I suggest later in this chapter and defend more fully in Chapter III, the rationale for the former type of delegation, on the content of decisions, does not necessarily justify the latter kind, on the principles and processes of decision-making in a democratic society. But, regardless of where political authority ultimately lies, we may expect practical confusion within individual institutions about what it means to make a legitimate decision at arm’s length until this void is addressed.

2. The historical context of the quango debate

The current ambiguity concerning what public decision-making at arm’s length ought to look like has a long and complex heritage, parts of which can shed light on the nature of the practical problems facing arm’s length governance today. Neither arm’s length bodies nor the controversies that surround them are entirely new in the UK: their roots may be traced back to 17th century quasi-autonomous boards,³⁵ like the Board of

³⁵ Carsten Greve, Matthew Flinders, and Sandra Van Thiel, “Quangos—What’s in a Name? Defining Quangos from a Comparative Perspective,” *Governance* 12, no. 2 (1999): 130.

Trade, a body primarily responsible for advising the British government in its mercantile affairs in colonial America.³⁶ And, in the wake of a boom in the creation of arm's length bodies, various attempts to classify, review, and reform these institutions have been made since the 1980s. While recognizing this broad sense of continuity, it is a set of subtle changes in both the political practice and critical theory of arm's length governance that is most useful for understanding the problems facing arm's length bodies today.

Situating the current state of arm's length governance within its broader history, this section presents the three most consequential developments for the articulation of these problems. These developments are drawn not only from shifts in the political manipulation and practical operation of arm's length bodies, what might be described as the history of arm's length governance, but also from developments in the political theory of and contemporary commentary upon this use, what might be described as the historiography of arm's length governance. These three areas of development are 1) the growing rhetorical emphasis on accountability rather than just cost-cutting in quango reform, 2) the seemingly trivial but in fact consequential changes in the classification of the arm's length landscape, and 3) the increasingly sophisticated understanding of the effects on political legitimacy of setting decisions at arm's length.

A. The rhetorical shift

It would oversimplify the nuances of the quango debate to say that earlier rationales for reform cited only the high cost and inefficiency of arm's length bodies, while later rationales focus upon their lack of transparency and accountability. Yet there

³⁶ *Encyclopædia Britannica Online*, s. v. "Board of Trade," accessed February 19, 2014, <http://www.britannica.com/EBchecked/topic/601629/Board-of-Trade>.

has been a gradual transformation of the rhetorical emphasis in the debate on arm's length bodies since the 1979 election of Margaret Thatcher's Conservative government. Leading up to the general election in 1979, the Conservative party targeted quangos as a prime example of government waste and inefficiency. Post-election, in the fanciful words of one political scientist, Thatcher "stalk[ed] the quango" in an attempt to cut both public spending and inefficiencies arising from the tendency to use political patronage to fill positions within arm's length bodies.³⁷ Her government established quinquennial reviews of quangos, predecessor of the current Maude tests and triennial reviews.

Continued by Blair's Labour government with some updated guidance, these quinquennial reviews had three aims:

1. To ensure that NDPBs are subject to regular and challenging review, taking into full account the views and needs of their customers, staff, and other stakeholders;
2. To ensure that the right structures for these bodies are put in place, so that they *deliver the Government's agenda effectively, and work in an efficient and integrated way*; and,
3. To provide a strong focus on *improving their future performance*, including the scope for more partnership working, and better use of technology (emphases added).³⁸

As these aims suggest, the main focus of arm's length body reform under Thatcher and her successor, John Major, was on ensuring more efficient and effective performance by NDPBs. Notably absent from these three central aims are the now-ubiquitous references to the accountability, transparency, or openness of quangos. As Flinders and two other experts on UK quango reform, Katharine Dommett and Katherine Tonkiss, argue, these two consecutive Conservative governments, while maintaining an anti-quango platform,

³⁷Robert Presthus, "Mrs. Thatcher Stalks the Quango: A Note on Patronage and Justice in Britain," *Public Administration Review* 41, no. 3 (1981): 312.

³⁸Cabinet Office, *Executive Non-Departmental Public Bodies 1999* (Cm 46, 1999): 22 cited in Katharine Dommett, "Triennial Reviews and Public Body Reform in the UK," ESRC Shrinking the State Project, January 2014.

only succeeded in the “cosmetic pruning” of the supposed overgrowth of arm’s length bodies, in part due to this limited scope of reform.³⁹

Taking up the ‘quango problem,’ the New Labour governments of Blair and Gordon Brown raised the profile of openness and transparency as qualities essential to the legitimate operation of arm’s length bodies. Although New Labour continued the quinquennial reviews established by Thatcher until 2003, it also published a set of reforms dedicated specifically to openness and transparency in *Quangos: Opening the Doors*.⁴⁰ A number of these reforms are still active today, including the practice of annual open public meetings, consultation with users via questionnaires or meetings, and consultation with members of the public on ‘matters of public concern.’

The Coalition government built upon this direction taken by New Labour. While one of the aims of the Maude review was undoubtedly to reduce the number and cost of arm’s length bodies, the Maude tests’ defining and unique feature is their description of the functions that the Government thinks are *best* carried out at arm’s length from elected politicians. In effect, this has carved out a particular space for arm’s length bodies, with the new question being *how* rather than *whether* such bodies should continue to operate. This is in contrast to the more extreme anti-quango rhetoric of the Conservative party under Thatcher, who was generally loath to admit the value of setting decisions at arm’s length. As I suggested in Section 1, a key shortcoming of the Coalition government was its failure to tackle the *internal* operation of quangos. Nevertheless, it attempted to strengthen the external lines of accountability between NDPBs and their sponsor departments, and explicitly brought many functions back

³⁹ Matthew Flinders, Katharine Dommett, and Katherine Tonkiss, “Bonfires and Barbecues: Coalition Governance and the Politics of Quango Reform,” *Contemporary British History* 28, no. 1 (2014): 4.

⁴⁰ “Government Opens the Doors on Quangos,” *Local Government Chronicle*, June 30, 1998, <http://www.lgcplus.com/government-opens-the-doors-on-quangos/1457666.article>.

within ‘democratically accountable’ departments, a focus absent under Cameron’s conservative predecessors due to their concentration on effectiveness and efficiency.

There has also been a new emphasis on the empty promise of the quango ‘numbers game’ since the election of the Coalition government. The Institute for Government in *Read before Burning*, one of the central reports on the most recent reform of arm’s length governance, stresses that “a simple cull of ALB [arm’s length body] numbers does not always amount to major reductions in cost” as most money is spent by a relatively small number of bodies.⁴¹ Similarly, Flinders and professor of public governance Chris Skelcher note, “The simple fact...is that abolishing quangos without abolishing functions does not save money.”⁴² Even Cameron in his election campaign shied away from the extreme ‘quango hunt’ called for thirty years before: “It would be far too simplistic for me to stand here and announce some kind of ‘Bonfire of the Quangos.’ People have heard that kind of talk many times before, and seen little to show for it.”⁴³ While the number and cost of quangos are still important figures, they are no longer the primary considerations in thinking about the operation of arm’s length bodies.

Indeed, the question of the accountability of arm’s length bodies has become central not just to the ‘quango debate,’ but to the broader preservation of a well-ordered democracy. In what remains his most detailed public address on arm’s length bodies, *People Power: Reforming Quangos*, Cameron attributes “the damaged trust in our political system” to accountability problems:

In a healthy democracy, the contract with the voters is simple. I voted you in. You’re responsible for what happens. If things go wrong, I’m going to make you answer for it. And if I don’t like the answer, I’m going to vote

⁴¹ Gash et al., “Read before Burning: Arm’s Length Government for a New Administration,” 10.

⁴² Matthew Flinders and Chris Skelcher, “Shrinking the Quango State: Five Challenges in Reforming Quangos,” *Public Money & Management* 32, no. 5 (2012): 332.

⁴³ David Cameron, “People Power: Reforming Quangos,” Speech to Reform, (July 6, 2009), <http://conservative-speeches.sayit.mysociety.org/speech/601336>.

you out...The problem today is that too much of what government does is actually done by people that no one can vote out, by organizations that feel no pressure to answer for what happens and in a way that is relatively unaccountable...*Any serious programme aimed at redistributing power from the political elite must address the role of quangos in creating this accountability problem and must include a serious plan to reform them* (emphasis added).⁴⁴

As this statement highlights, the increasing delegation of government functions to unelected bodies means that the potency of traditional democratic voting as a method of ensuring state accountability is diluted: voters cannot exercise the same control over quangos that they do with respect to elected politicians. Arm's length bodies may be subject to indirect democratic control to the extent that voters view their MPs as ultimately responsible for delegated decisions and reward or punish them in elections accordingly, which will incentivise politicians to control quangos in line with their constituents' desires. But this indirect effect is likely to be relatively weak and take longer to impact governance, given the number of other considerations upon which citizens vote for their representatives. Furthermore, if we think there are in fact good reasons to place some decisions beyond the reach of democratically elected officials, this kind of control will defeat the purpose of setting them at arm's length.

As people increasingly come to the conclusion that 'quango counting' or reference to the total sum of public funds spent by quangos is a non-starter, questions of accountability and transparency in arm's length governance have risen in public profile. Any practical and rigorous theory of what governance at arm's length ought to look like will require an understanding of how these concepts interact with the independence from elected officials that arm's length bodies are designed to ensure. How can bodies designed to make politically impartial decisions also be held accountable? How can frequently controversial expert judgments, those made by "people with the right training,

⁴⁴ Ibid.

professional knowledge and specialist skills,”⁴⁵ in Cameron’s words, be made suitably transparent to the public? I return to these questions in Subsection C below.

B. Tweaking the taxonomy

A second development over the past thirty years is the way in which arm’s length bodies are classified. Due to the sheer variety and number of bodies, there have been many attempts to account for and classify the entire spectrum. On the surface, such taxonomies may appear to be of merely administrative interest. However, the successive attempts at reclassification often mask, for political purposes, important similarities between supposedly distinct classes. This insight is important for the groundwork of this dissertation as it explains why I consider the MHRA, an executive agency frequently classed out of the quango landscape, as an arm’s length body comparable to the HFEA and NICE, two executive NDPBs, the traditional archetype of the quango.

The acronym ‘quango’ was coined by British academic Anthony Barker in the early 1970s for a term first introduced by Alan Pifer, then president of the Carnegie Corporation of New York, in a 1967 essay, “The Quasi Nongovernmental Organization.”⁴⁶ Pifer used the term to describe “the difficulty of reconciling the needs for independence and accountability in a new type of enterprise – incorporated in the private sector but funded by the Federal Government – that had begun to appear.”⁴⁷ Predicting, ahead of his time, the important tension between accountability and independence in arm’s length bodies, Pifer had in mind not-for-profit corporations providing advice to various branches of the US Armed Forces, technical assistance

⁴⁵ Ibid.

⁴⁶ Alan Pifer, “Letter: On Quasi-Public Organizations; Whence Came the Quango, and Why,” *The New York Times*, September 5, 1987, sec. Opinion, <http://www.nytimes.com/1987/09/05/opinion/1-letter-on-quasi-public-organizations-whence-came-the-quango-and-why-969587.html>.

⁴⁷ Alan Pifer, “The Quasi Nongovernmental Organization,” *Annual Report for 1967* (Carnegie Corporation of New York, 1967), <http://files.eric.ed.gov/fulltext/ED018846.pdf>, 4.

programs sponsored by the Agency for International Development, and community action agencies supported by the Office of Economic Opportunity. While the origins of the term thus trace to the United States, Barker's popular abbreviation anticipated the distinctly British path this institutional class would follow. Indeed, while there are certainly arm's length bodies in the US, 'quango' has not been utilized – as either a formal or informal term - in the classification of the American political system. As such, while 'quango' has become a household term in the UK, it is largely unfamiliar in the US.

However, the types of bodies classified as quangos in the UK have varied over time. Under Thatcher and soon after Barker's introduction of the term, civil servant Leo Pliatzky conducted a review of quangos in 1980 that was distinctive "for the manner in which it linked the term 'quango' with just one specific form of arm's-length body – the Non-Departmental Public Body (NDPB)."⁴⁸ By excluding operationally similar bodies, such as executive agencies, Pliatzky was able to report a smaller number and total cost of quangos. This kind of taxonomical manipulation has continued to the present day, as successive governments and opposition parties have attempted to doctor the appearance of the quango landscape for political purposes. This has made it difficult for any two assessors to come to the same conclusion about which organizations are quangos, which are not, and why, resulting in widespread uncertainty about the status and proper function of such bodies. Writing in the aftermath of the Labour government's election in 1997, Flinders and two other political theorists note the continuity of this trend, characterizing it as a case of "institutional and terminological confusion [within British quasi-government] which are both fundamentally linked."⁴⁹ And over a decade later,

⁴⁸ Flinders, Dommett, and Tonkiss, "Bonfires and Barbecues," 8.

⁴⁹ Greve, Flinders, and Van Thiel, "Quangos—What's in a Name?," 137.

suggesting the persistence of confusion, the Public Administration Select Committee echoed this sentiment in its response to the Maude tests and the Coalition government's Public Bodies Reform program:

The current system is chaotic, making it difficult to understand why different types of bodies exist and what these variations mean in practice. We recommend that the government use its triennial review process to re-examine the proper governance arrangements for each public body and place them in a new simplified taxonomy.⁵⁰

Indeed, when the Coalition government set out to map the existing landscape of arm's length bodies, Maude admitted, "It may sound absurd but it was and remains incredibly difficult to gain firm information on such bodies. Many do not publish accounts, there is no central list, and there are myriad different types all with different statuses."⁵¹

While, as suggested by Flinders and Skelcher, "the 20th century was littered with official committees, parliamentary inquiries and academic projects that attempted to map out the topography of the quango state,"⁵² the most recent and arguably most valuable taxonomical endeavour is that of the Institute for Government's *Read Before Burning*. Research for the report took place nearly concurrently with Maude's review in 2010. Tellingly, while the Government's own classification system identified three types of bodies – NDPBs, non-ministerial departments, and public corporations – the Institute's review of the arm's length landscape presented a competing account that identified no less than 11 types: 1) advisory NDBPs, 2) executive NDPBs, 3) independent monitoring boards, 4) tribunal NDPBs, 5) executive agencies, 6) non-ministerial departments, 7) public corporations, 8) independent statutory bodies, 9) special health authorities, 10)

⁵⁰ Cabinet Office, *Government Response to the Public Administration Select Committee Report, "Smaller Government: Shrinking the Quango State,"* 26.

⁵¹ HC Deb, 14 October 2010, vol 516, col 508.

⁵² Flinders and Skelcher, "Shrinking the Quango State," 330.

parliamentary bodies, and 11) the central bank.⁵³ Perhaps the most glaring discrepancy between the two was the failure of the Government to include executive agencies in its review. This decision, according to criticism from Flinders and Skelcher, was based on “a constitutionally dubious distinction between executive agencies and executive NDPBs.”⁵⁴

The Government, while recognizing the importance of accountability in the assessment of arm’s length governance, as suggested in Subsection A, nevertheless perpetuated the political habit of using reclassification to mask the shortcomings of actual reform. Rather than tackling the admittedly complex question of how to preserve arm’s length bodies’ necessary independence while also guaranteeing their accountability, the Government created a black-and-white distinction between executive agencies as a class, ‘accountable’ because they are hived into a ministerial department, and NDPBs as a class, ‘independent’ because they are hived out. After making this distinction, many NDPBs were simply reclassified as or within executive agencies. The Government, whether naively or deceptively, suggested that by changing little more than the superficial location of bodies on an organizational chart, a large portion of the quango problem was ‘solved.’

The Institute for Government, on the other hand, created a new, simplified taxonomy with the goal of “enabl[ing] form to be directly related to function,” where “the classification is determined by the degree of freedom from executive control on appointments, strategy, decisions and budget, which the body needs to be able to discharge its functions.”⁵⁵ In other words, its taxonomy is meant to confront head-on the accountability-independence tension by making the degree of independence from

⁵³ Gash et al., “Read before Burning: Arm’s Length Government for a New Administration,” 22.

⁵⁴ Flinders and Skelcher, “Shrinking the Quango State,” 332.

⁵⁵ Gash et al., “Read before Burning: Arm’s Length Government for a New Administration.”

ministers the defining axis of the arm's length body spectrum. In place of the 11 classes it initially identified, the Institute suggested four new categories: 1) constitutional bodies, 2) independent public interest bodies, 3) departmental sponsored bodies, and 4) executive agencies, with constitutional bodies being the most independent from ministers and executive agencies being the closest to ministerial control.

Constitutional bodies, like the Electoral Commission, would be free from executive control and instead accountable to Parliament, as their independence is crucial to their constitutional oversight. Independent public interest bodies, which would include regulators, guarantors of standards, and independent watchdogs, require independence in order to secure the credibility of their work. As such, while ministers could have some influence on the policy direction of such bodies, parliamentary oversight via the relevant select committee would check this executive power. Departmental sponsored bodies carry out core functions of one or more departments, but can more efficiently carry out their roles with day-to-day independence from the department. Most of these bodies require special technical expertise, and include bodies that hand out grants, perform enforcement and inspection roles, or are responsible for stewardship of national assets. Finally, executive agencies would be defined as they are now, as independent business units that implement government policy but that are nevertheless constitutionally part of a department and thus held to account by their minister.⁵⁶

Theoretically, this taxonomy is a promising development: both classification and governance arrangements follow from the necessary independence needed by a given public body to perform its function. Furthermore, it includes executive agencies in the spectrum of arm's length governance, recognizing that an agency still operates independently from its parent department even if it is ultimately accountable to its

⁵⁶ Ibid., 55-57.

minister. This is an important recognition because poor oversight from parent departments can, as the Office for Public Sector Reform noted in 2002, lead agencies to assume “orphan status,” and thus ‘drift’ away from ministerial control and come to resemble NDPBs. This substantial independence and the corresponding potential for drift, combined with the suspect reclassification of many NDPBs as or within executive agencies by the Coalition government, is a key reason that I include the MHRA as an arm’s length body regulating access to emerging biotechnologies.

In spite of this strength, the Institute’s taxonomy is such a radical restructuring that it has not and is unlikely to be put fully into practice; it suggests, for instance, the complete erasure of the non-ministerial department category. Furthermore, it suggests that advisory NDPBs should be reconstituted as expert advisory committees and no longer classified as arm’s length bodies because they do not have an executive function: “Bodies providing advice to ministers with no independent budgets and staff would no longer be regarded as part of the arm’s length landscape...The independence of such bodies would be secured via [protocols on the status of the advice of expert advisory committees] and through the reputation and expertise of advisors.”⁵⁷ However, it is the degree of freedom from ministerial control required to carry out a body’s function that is the defining feature of the Institute’s taxonomy, not whether it has an independent budget or provides advice instead of executive decisions, so this omission is peculiar. Expert advisory committees are indeed unelected public bodies whose advice we want to be politically impartial, and as such they seem to clearly fall in the arm’s length landscape. The ‘reputation and expertise’ of advisors cannot guarantee independence when it is these very qualities that may be in question, as is often the case in regulation of emerging biotechnologies. Political theorist Alfred Moore argues this point using stem

⁵⁷ Ibid., 56.

cell research and climate change as examples of cases where “credibility, reliability and neutrality of expertise are at the heart of disputes, and it is therefore clear that in such cases expertise cannot easily function as a neutral ground on which to resolve disputes.”⁵⁸ Furthermore, simply hiving these advisory bodies into departments or executive agencies obscures the way in which such expert advice ultimately affects policy, particularly worrying in the regulation of biotechnologies as the role that scientific expertise ought to play is frequently contested. To the extent that recommendations from such advisory bodies are simply ‘rubber stamped’ by the relevant minister, they thus present many of the same problems as executive NDPBs that draw on technical expertise.

This section has considered merely a few of the many past attempts to order the puzzling spectrum of arm’s length governance. By examining both the causes of taxonomical changes, like politically motivated ‘quango counting,’ and the effects of these shifts, particularly the obscuration of executive agencies and advisory NDPBs, I have hoped to explain two related decisions I have made regarding the scope of this dissertation. First, I consider the degree of practical autonomy from democratically elected officials to be the main feature of arm’s length governance, as it is this degree of independence that creates the special tension with accountability. Based on this basic principle, I have also decided to consider the institutional form of the executive agency alongside the NDPB in my analysis of arm’s length bodies that regulate access to emerging biotechnologies, even though various other sources count the former outside the arm’s length body landscape; in practice, both may function similarly and face the same challenges to operating in an accountable manner.

⁵⁸ Alfred Moore, “Between Competence and Consent: Democratic Theory and Expertise,” in *Expertise and Democracy*, ed. Cathrine Holst, vol. 1, 14 vols., ARENA Report (Oslo: ARENA, 2014), 52.

The next and final of these three subsections will examine the relationship between the concepts of independence, accountability, transparency, and legitimacy in arm's length governance, crystallizing the central tensions already suggested in the past two sections.

C. Mapping conceptual relationships

In her 2002 BBC Reith Lectures, philosopher Onora O'Neill focused on the UK's recent quest for "more perfect accountability" and "more complete openness and transparency" in public life, attributing this quest to the supposed 'crisis of trust' that pollsters suggest is plaguing contemporary society.⁵⁹ Unconvinced of the value of traditional methods of improving accountability and increasing transparency, she argues that we need 'intelligent accountability,' which "requires more attention to good governance and fewer fantasies about total control," where good governance requires "some margin for self-governance of a form appropriate to [institutions'] particular tasks."⁶⁰ Likewise, she suggests that total transparency may do more harm than good: "Demands for universal transparency are likely to encourage the evasions, hypocrisies and half-truths that we usually refer to as 'political correctness,' but which might more forthrightly be called either 'self-censorship' or 'deception.'"⁶¹ Her lectures raise a sweeping set of questions. How much control and how much freedom should public bodies have if they need the former to be publicly accountable and the latter to carry out their 'particular tasks'? Will total transparency damage the integrity of certain decisions? Is balancing accountability and autonomy a zero-sum game in which a body that is more autonomous is necessarily less accountable?

⁵⁹ Onora O'Neill, *A Question of Trust: The BBC Reith Lectures 2002* (Cambridge: Cambridge University Press, 2002), 45, 68.

⁶⁰ *Ibid.*, 58.

⁶¹ *Ibid.*, 73.

While her lectures apply to the broader public sector, the quango debate may serve as the crucible in which the proper relationships between these concepts are cast. As this dissertation seeks to give an account of how arm's length bodies should regulate emerging biotechnologies – understood as a specified question about the circumstances under which arm's length governance in a democratic society is legitimate – the way in which autonomy, accountability, and transparency either shore up or undermine legitimate decision-making is of central importance. In this section, I begin by examining how the relationships between these complex concepts have been mapped in the past. In particular, I focus on how an *oversimplified* modelling of these relationships has plagued previous thinking about arm's length governance. This shortcoming reveals the kinds of questions that must be addressed in formulating a legitimate conception of arm's length governance of emerging biotechnologies, which I delineate in the final section of this chapter.

Recall from Section 1 that the Maude tests were created and applied in the 2010-2011 review of public bodies “with the aim of increasing accountability for actions carried out on behalf of the state”:

This government made a presumption that state activity, if needed at all, should be undertaken by bodies that are democratically accountable at either national or local level. A body should only exist as a quango if it meets one of three tests... 1) Does it perform a technical function? 2) Do its activities require political impartiality? 3) Does it need to act independently to establish facts?⁶²

In this formulation, quangos are the *exception* to the standard presumption of democratic accountability, and the three tests set out the exceptional functions that ought to be performed at arm's length from democratically elected officials. The Coalition government's approach to reform consisted largely in abolishing quangos and

⁶² HC Deb, 14 October 2010, vol 516, col 27WS.

reassigning their functions to more democratically accountable bodies, like departments. Although this appears to address the “aim of increasing accountability for actions carried out on behalf of the state,” it only recognizes and addresses half of the problem. Bodies that remained at arm’s length after the review and cull received little to no reform or guidance. The Coalition government therefore managed to avoid the difficult question of what accountability *within* arm’s length governance should look like.

The Cabinet Office has made some limited attempts to fill this void in the wake of the initial review. It refers quangos that survived the review to its ‘Principles of Good Corporate Governance’ for NDPBs, where it suggests three principles of accountability: 1) the principle of statutory accountability, requiring arm’s length bodies to comply with statutes, regulations, and statements of best practice, 2) the principle of accountability for public money, requiring them to properly report to Parliament on the use of public money, and 3) the principle of ministerial accountability, which – while requiring nothing from quangos directly – simply states, “The Minister [of the sponsoring department] is ultimately accountable to Parliament and the public for the overall performance of the public body.”⁶³ In line with the third principle, the Government has attempted to revitalize the role of departmental sponsorship, publishing in April 2014 an *Introduction to Sponsorship* guide for the Civil Service.⁶⁴ Both the principles of good corporate governance and the sponsorship guide focus on maintaining clear lines of communication so that both the department and the arm’s length body are on the same page. While an important step, it provides little practical help in determining – once communication is adequately clear – how much independence the arm’s length body

⁶³ Cabinet Office, “Categories of Public Bodies: A Guide for Departments,” Annex B.

⁶⁴ Civil Service, “Introduction to Sponsorship: An Induction Pack for New Sponsors of Arm’s-Length Bodies” (London: Civil Service, April 2014), https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/303583/Introduction_to_Sponsorship_An_Induction_Handbook_for_New_Sponsors_of_Arm_s-Length_Bodies.pdf.

should have in its operations. It is a ‘soft’ accountability, in which bodies must transparently keep their departments abreast of what they are doing, rather than a ‘hard’ accountability, which would define the length of the ‘leash’ or degree of freedom such bodies are granted from ministers in decision-making. This soft version of accountability-as-transparency does not help in determining, for instance, on what grounds a department can or should overrule the decision of an arm’s length body or, by a less direct means of influence, reduce its funding.

One response to this lingering question of how arm’s length bodies should be held accountable is that they simply shouldn’t be: they are purposefully set up to be independent and thus discussion of their accountability misses the point. This view is reflected in the research of the Carnegie UK Trust’s Local Government Information Unit (LGIU), in which one interviewed minister said that most government departments do not manage their public bodies because “the whole point is these are meant to be autonomous and not accountable, so if there’s a justification for the function being carried out in a way that’s independent of a department, then the ability of the sponsoring department to interfere with its management is strictly limited.”⁶⁵ The tension between independence and accountability is rendered particularly clear in this statement: on the one hand, we want certain decisions to be politically impartial (thus the need for independence from elected politicians). On the other hand, we want to ensure that such public decisions are made in the public interest (thus the need for some form of accountability). This minister essentially represents the view that we can’t have it both ways. Two points are worth noting about this ‘paradox position.’ First, this interpretation of the relationship between independence and accountability assumes an all-or-nothing,

⁶⁵ Laurie Thraves, “The Future of Arm’s-Length Bodies” (London: Local Government Information Unit, March 19, 2012), <http://www.lgiu.org.uk/wp-content/uploads/2012/04/Future-of-arms-length-bodies.pdf>, 14.

zero-sum game: a public body is either autonomous or accountable but not both. Second, it assumes that what public bodies are ‘independent from’ or ‘accountable to’ are necessarily the same entity, in this case the department: a public body is either independent from a department or accountable to it.

The assumptions of the paradox position about autonomy and accountability can both be challenged. Greve, Flinders and van Thiel note that while autonomy in the quango debate is usually conceived of as “autonomy from elected politicians (or the state)”, the ‘autonomy from whom’ question can be answered in a variety of ways based on the rationale for the arm’s length body’s creation.⁶⁶ For instance, certain bodies may require independence from market forces or powerful interest groups rather than their department and minister.⁶⁷ Likewise, accountability need not be conceived of in such a narrow way. Thinking of accountability as only ‘accountability to elected officials’ “ignores the wider sense of the term, which involves being accountable to the public forums of civil society.”⁶⁸ Combining these two broadened accounts of independence and accountability allows for the possibility, precluded in the paradox position, that arm’s length bodies can be *both* independent and accountable. In decisions where political impartiality is needed, for instance, we might ideally want a governance structure in which the body responsible for the decision is autonomous from its department and minister but accountable to the public, where ‘the public’ is more broadly conceived of than those democratically elected officials representing the current majority.

After describing the paradox position in which ministers seem to find themselves, the LGiU suggests labelling arm’s length bodies as either ‘independent’ or ‘politically

⁶⁶ Greve, Flinders, and Van Thiel, “Quangos—What’s in a Name?,” 140.

⁶⁷ Ibid.

⁶⁸ Thraves, “The Future of Arm’s-Length Bodies,” 6.

accountable.’ Bodies labelled independent would be “those whose primary purpose is to provide independence of decision-making or advice,” while all other bodies should be considered politically accountable, with a clearly identified sponsoring politician who provides an annual ‘letter of mandate’ describing the function that the arm’s length body must carry out.⁶⁹ Such labelling would usefully distinguish between bodies that are set up at arm’s length to provide politically impartial decisions and those set up to reap the efficiency gains of day-to-day autonomous operation. However, the LGiU provides no accountability framework or guidance for the operation of those arm’s length bodies labelled ‘independent,’ reinforcing the paradox position and its narrow concept of accountability. Instead, it only suggests a need for greater transparency in such independent bodies, noting simply, “Transparency in governance is an important source of political legitimacy.”⁷⁰ Again, while important, it is far from clear how transparency is meant to bolster legitimacy if there is no defined concept of the principles upon which such independent bodies should operate and no method of holding them to account for that operation. In other words, even if we can clearly *see* how a body operates, this will be of limited use if there is no public conception of how it should operate or any means to hold it to this account.

Recognizing the simplistic sense in which accountability has been conceived in the past, recent theories of arm’s length governance have attempted to reconceptualise accountability in the arm’s length context, working towards O’Neill’s notion of ‘intelligent accountability.’ In his influential book on arm’s length governance, Flinders calls for scholars to:

...[M]ove away from and reject implicitly binary classifications which seek to label organizational forms as either accountable/unaccountable,

⁶⁹ Ibid., 15

⁷⁰ Ibid., 16.

legitimate/illegitimate, command/control, private/public, democratic/undemocratic – terms that are as crude as they are blunt – and instead seek to tease apart the specific meanings and connotations of these terms.⁷¹

With respect to legitimacy, he notes that ‘electoral’ legitimacy is just one of many important types; legitimacy may also be understood in virtue of expertise, experience, or objectivity.⁷² These different accounts of legitimacy implicate radically different concepts of accountability. A decision by experts, by scientists for example, may be legitimate in virtue of its epistemic quality – accountable, in some sense, to the scientific truth of the matter – but will not seek to be democratically accountable. A decision heavily informed by those with specific experience of the situation in question, perhaps a genetic condition, may be legitimate in virtue of this lived experience – accountable to those affected by it and the way in which they understand their condition – but will not seek to be accountable to the general public or its beliefs about this condition. These various accounts of the relationship between accountability and legitimacy contribute to what Flinders calls the “arguably incompatible, even antagonist” models of legitimacy in British governance.⁷³

Questions of legitimacy also require better scrutiny of the situations in which arm’s length governance is appropriate in the first place. While the thrust of this analysis has been thus far on how arm’s length bodies that pass one of Maude’s three tests ought to operate, a preceding question is whether these tests of technical expertise, political impartiality, and independence in establishing facts are in fact the right ones. In other words, we must first decide the contexts in which independence from elected politicians is legitimate before determining what that autonomous operation should look like. The

⁷¹ Flinders, *Delegated Governance and the British State*, 309-310.

⁷² *Ibid.*, 319.

⁷³ *Ibid.*, 320.

three tests devised by the Cabinet Office seem to be merely *post hoc* rationalizations of the types of functions that already existed at arm's length. As the authors of *Read Before Burning* note, the Cameron government's public bodies review thus "represents a missed opportunity to open up a debate about which functions are best performed at arm's length and which are not."⁷⁴ All seem to agree that there are at least some functions that ought to be performed at arm's length, for instance, auditing of campaign financing and elections, and inquiries into unfair treatment of individuals by government departments. Whether and how such arm's length governance should be used for decision-making on other politically, morally, and scientifically controversial issues, like access to emerging biotechnologies, is far less clear.

3. The legitimacy problem in the arm's length context

This chapter has examined the historical context and current state of arm's length governance in the UK in order to tease out the central issues of the perennial 'quango problem.' In this final section, I step back from the history and current politics of quangos to provide a decontextualized account of the legitimacy problem for decision-making at arm's length. This shared theoretical problem manifests itself in practical challenges and confusion within the otherwise quite distinct contexts of the HFEA, NICE and MHRA; as such, any attempt – like mine – to give an account of how we ought to make collective and binding decisions about public access to emerging biotechnologies must recognize and confront this problem head-on.

In a well-ordered democratic society, the normative legitimacy of government action is usually ensured by free and fair elections: laws and regulations are tied to the politicians who pass or construct them, and those politicians are held accountable for

⁷⁴ Gash and Rutter, "The Quango Conundrum," 99.

their decision-making by the people who either re-elect or vote them out. Nevertheless, there are certain functions that most can agree are best carried out at arm's length from elected politicians. While the current Government maintains that these are functions that require technical expertise, political impartiality, or independence in establishing facts, this list is certainly open to challenge.

The quangos performing these functions make decisions on behalf of the public about the way citizens can and cannot lead their lives. In the case of biotechnology regulation, this includes decisions about whether and when the state should “interfere with the liberty of citizens to make use of the opportunities offered to them by modern science and medicine,”⁷⁵ whether, for instance, citizens should be permitted to ‘design’ their children, receive promising but expensive new treatments on the national healthcare system, or utilize novel drugs and devices. Arm's length bodies thus ultimately wield significant coercive power on behalf of the state in regulating what individuals can and cannot do.

Because these arm's length bodies have been formed with the express purpose of removing certain decisions from elected politicians, the traditional notion of electoral accountability is inadequate for ensuring that bodies set these coercive regulations in a legitimate manner, one that can be publicly justified in a democratic society. In order to operate ‘in the public interest’ or on behalf of ‘the people’ broadly conceived, arm's length bodies thus need a way to think about making legitimate, binding collective decisions in the absence of elected representatives. To date, most accounts of legitimacy and accountability in delegated governance focus narrowly on making the process of decision-making and the structure of the institution more ‘transparent,’ abstaining from

⁷⁵ Silke Schicktanz and Michael Dusché, “The Ethics of Ethical Expertise in Science, Medicine and Healthcare Policies,” *Indian Journal of Medical Ethics* 8, no. 3 (2011): 142.

any more substantial principles about how decisions should be made. How can we conceptualize a ‘thicker’ account of normative legitimacy in the context of such delegated decision-making? And how can these bodies be held publicly accountable to this conception?

Some have suggested, as implied by the ‘paradox position,’ that we must simply give up on any externally determined conception of legitimacy or accountability when we place decisions at arm’s length, leaving each institution to settle these conceptions for itself as part and parcel of its delegated function. As I will argue in Chapter III, while I think there are good arguments for keeping decisions about access to emerging biotechnologies at arm’s length, I do not think that this delegation requires the concomitant and complete deferral of authority entailed by the paradox position. Rather, I argue that such bodies should be guided by a basic conception of public accountability – rooted in democratic principles of public reasoning but quite distinct from traditional electoral accountability – and a framework of process-oriented conditions for decision-making that follow from this conception.

However, it is worth realizing that the practical challenges for arm’s length bodies that I illustrate in the following chapter exist regardless of where authority for determining what legitimate arm’s length decision-making looks like. The HFEA, NICE, and MHRA each struggle with the practical manifestations of this legitimacy problem within their unique settings, without necessarily recognizing how these challenges are embedded within their shared politico-institutional form. They may be narrowly focused on questions of how public consultation should feed into decision-making, or what is required to be properly responsive to the patient voice. But these practical challenges will persist until an account of legitimate arm’s length decision-making is articulated that can justify a role for these practices, regardless of whether that account is set by

Parliament or by each individual body myopically ‘muddling through’ this problem. While I argue that this account can and should be set at the state level, and then fleshed out within the three different settings, merely rejecting this location of authority in favour of the paradox position does not resolve the very real, practical challenges before each body. It simply alleges that there are *no* general substantive or procedural principles that can guide these decisions and that each arm’s length body must be left to its own devices to work it out.

This approach is, in the first instance, unjustifiable, as these decisions substantially impact citizens’ lives and should be constrained by some conception of public accountability to ensure that they are in fact made in the public interest. And, in the second instance, it is unhelpful, given that it merely passes the responsibility for addressing the legitimacy problem along to each individual arm’s length body, greatly duplicating work on a common problem and contributing to inconsistency in how this problem is addressed across bodies that regulate various types and stages of emerging biotechnologies. Chapter III provides my account of how we ought to navigate this shared legitimacy problem and a new framework for public accountability in the arm’s length context. But first, in the following chapter, I introduce each of the three bodies that I have touched upon briefly thus far and illustrate how the general theoretical problem articulated here manifests itself practically within each body.

Chapter II. Key Players and Practical Challenges

The Human Fertilisation and Embryology Authority (HFEA), the National Institute for Health and Care Excellence (NICE), and the Medicines and Healthcare Products Regulatory Agency (MHRA) are the three main sites at which the permissibility of public access to emerging biotechnologies in the UK is determined. To be sure, there are many other organisations that contribute to these determinations at different temporal stages and from different institutional locations. For instance, the impact of the Nuffield Council on Bioethics is widely recognized and its reports – like those on donor conception⁷⁶ and novel neurotechnologies⁷⁷ – are frequently cited in policy documents. However, the Council is an independent, non-governmental body (arguably a large factor in its impact) with no decision-making powers; while it provides useful and important advice to policymakers, it is ultimately the responsibility of members of the HFEA to set regulations for donor conception and that of members of the MHRA to determine whether to license new neurotechnologies.

On the other hand, the Health Research Authority is a government body, a non-departmental public body (NDPB) like the HFEA and NICE, that sets standards and guidelines for research ethics and has recently begun approving research itself. Its decisions about the proper balance of risks and benefits in health research or the protection of privacy in trial design may certainly affect which new drugs and devices

⁷⁶ Nuffield Council on Bioethics, “Donor Conception: Ethical Aspects of Information Sharing” (London: Nuffield Council on Bioethics, April 2013), <http://nuffieldbioethics.org/project/donor-conception/>.

⁷⁷ Nuffield Council on Bioethics, “Novel Neurotechnology: Intervening in the Brain” (London: Nuffield Council on Bioethics, June 2013), <http://nuffieldbioethics.org/project/neurotechnology/>.

are ultimately available for public consumption, as well as how quickly they reach the market. But this effect on public access is further ‘upstream’ than those of the three bodies considered here, which must decide after such research whether to license, in the case of the MHRA, or publicly provide, in the case of NICE, the new product.⁷⁸ While, as an arm’s length body, the decision-making of the Health Research Authority may be subject to and benefit from the public accountability conditions I set out in the next chapter, it is not a direct gatekeeper to emerging biotechnologies and thus falls beyond my scope.

As a final example of an institution that I have not chosen to consider in depth, the National Health Service (NHS),⁷⁹ another executive NDPB, is the most direct point of access to new drugs, devices, and procedures for citizens. However, as I discuss at greater length in Chapters V and VI, an individual doctor’s decision about what treatment is most appropriate for an individual patient is distinct from, and constrained by, national licensing decisions from the HFEA and MHRA and resource allocation decisions from NICE. The individual consultation relies upon a doctor-patient conversation about the patient’s specific values, symptoms, and biological characteristics in order to choose amongst a range of options, while national licensing and allocation decisions require additional attention to *collective* considerations, like the total healthcare budget or public health and safety, as well the existence of pluralistic values on controversial questions like reproduction. While the former may be understood largely as only affecting the patient, the latter decisions may impact others beyond the

⁷⁸ Furthermore, the HFEA itself is responsible for approving, through its licensing scheme, research on human embryos.

⁷⁹ National healthcare is provided by different bodies across the UK: NHS England, NHS Scotland, NHS Wales, and Health and Social Care in Northern Ireland. Formally, legislation established NICE as providing guidance only to England. However, it informally provides certain types of guidance, including technology appraisals, to the devolved governments. In contrast, the HFEA and MHRA are formally UK-wide regulators.

specific ‘user’ and thus must take these collective considerations into account. As such, these national licensing and allocation decisions must be made in a publicly accountable manner, but subsequent decisions about what – of those licensed or funded treatments – should be used by an individual patient are properly the private and collaborative decisions of the patient and his doctor. In sum, these later decisions are not the type of *public* decisions made by the three bodies I consider, and thus also fall out of the scope of this dissertation.

Having briefly explained my identification of the HFEA, NICE, and MHRA as the primary bodies regulating public access to emerging biotechnologies, I turn now to the chapter’s main goal of describing the remit of each institution and how the legitimacy problem manifests itself in each arm’s length setting. Each section gives a brief history of the body, what it regulates, and its current structures and principles for making these regulatory decisions. Then, drawing on a range of sources, including policy documents, parliamentary and departmental oversight reports, and both my own and others’ past empirical research, I illustrate a central challenge to decision-making that each body is currently struggling to address. In providing this account, I demonstrate how each challenge is implicated by the institutional situation at arm’s length and the tension it creates between independent and accountable decision-making. After articulating a solution for this shared legitimacy problem in Chapter III, I return to these three practical manifestations of the problem in Chapters IV, V, and VI, where they serve as case studies for how my public accountability framework can be practically and fruitfully applied within various settings.

1. Human Fertilisation and Embryology Authority

After the 1978 birth of Louise Brown, the first child conceived through the use of *in vitro* fertilization (IVF), the Department of Health assembled a Committee of Inquiry into Human Fertilisation and Embryology in 1982 “to examine the social, ethical, and legal implications of recent, and potential developments in the field of human assisted reproduction.”⁸⁰ More commonly referred to as the Warnock Committee for its chair, Oxford philosopher Mary Warnock, the Committee released its report two years later. It noted the inability of existing professional codes to sufficiently regulate the recent developments in IVF and recommended, in addition to a number of protections for the human embryo based on its “special status,” the creation of a statutory body charged with regulating and licensing research and treatment.⁸¹

After a number of failed attempts in Parliament to block all research on human embryos, the *Human Fertilisation and Embryology Act 1990* (HFE Act 1990) was passed. Following the Warnock Committee’s recommendations, the Act explicitly banned reproductive cloning, embryo research after the appearance of the ‘primitive streak’ (approximately 14 days after fertilisation), and the implantation of embryos with altered nuclear or mitochondrial DNA. The regulation of all other, non-prohibited activities, including the creation of embryos for use in treatment and research, the use of donated gametes and embryos, and the storage of gametes and embryos, was delegated to the HFEA, created as an executive NDPB.⁸² The next major parliamentary review of

⁸⁰ “Report of the Committee of Inquiry into Human Fertilisation and Embryology” (Cmnd 9314, July 1984), http://www.hfea.gov.uk/docs/Warnock_Report_of_the_Committee_of_Inquiry_into_Human_Fertilisation_and_Embryology_1984.pdf, 4. For a detailed history of the creation HFEA, see Wilson, “Creating the ‘Ethics Industry.’”

⁸¹ Ruth Deech and Anna Smajdor, *From IVF to Immortality: Controversy in the Era of Reproductive Technology* (Oxford: Oxford University Press, 2008): 8.

⁸² *The Human Fertilisation and Embryology Act of 1990*, c. 37, <http://www.legislation.gov.uk/ukpga/1990/37/contents>.

the HFE Act 1990, and the only to date, began in 2004.⁸³ The review ensured that the law was up to date with technological advances that had occurred since 1990, and brought a number of the HFEA's policies into law. Its main changes included a ban on sex selection for non-medical reasons, removal of the 'need for a father' in assessing "the welfare of the child" before providing fertility treatment, recognition of same-sex partners as legal parents of a child conceived with a donor embryo or donor gametes, and clarification of rules for hybrid embryo research.⁸⁴ These new provisions and amendments to the HFE Act 1990 were passed in 2008. Most recently, Parliament voted in February 2015 to amend the total ban on altering the DNA of embryos used in treatment, in order to permit the use of mitochondrial replacement techniques to avoid serious mitochondrial diseases. However, implanting embryos with altered nuclear DNA is still legally prohibited.

A. Decision-making principles and processes

The core of the HFEA is a small decision-making committee seating only 12 members as of August 2015, who are appointed for three-year terms with an opportunity for reappointment. Based on recommendations from the Warnock committee about the importance of lay membership, by law both the chairman and deputy chairman must be 'lay and unbiased.' This requirement excludes any person who is or has been 1) a registered medical practitioner, 2) concerned with keeping or using gametes or embryos outside the body, or 3) concerned with commissioning, funding, or participating in

⁸³ Department of Health, "Review of the Human Fertilisation and Embryology Act: Proposals for Revised Legislation (including Establishment of the Regulatory Authority for Tissue and Embryos)" (Cm 6989, December 2006), http://www.hfea.gov.uk/docs/Review_HFEA_Act_White_Paper_DH.pdf.

⁸⁴ *The Human Fertilisation and Embryology Act of 2008*, c. 22, <http://www.legislation.gov.uk/ukpga/2008/22/contents>.

research on keeping or using gametes or embryos outside the body.⁸⁵ Additionally, at least one-third but less than half of the members must meet one of these three requirements, and one ‘super-lay’ member must meet all three.⁸⁶ Thus, in addition to scientific experts like geneticists, molecular biologists, clinical embryologists, and fertility clinicians, members of the HFEA have included a broad range of individuals, such as philosophers, lawyers, consultants, journalists, and bishops.

Members of the HFEA assemble all together at least six times per year to discuss major issues and new policy proposals at meetings that are open to the general public. However, members are also involved in a number of sub-committees that make more routine decisions on behalf of the HFEA. For instance, the License Committee grants and renews general licenses for treatment and research to individual fertility centres, the Statutory Approvals Committee considers applications to use pre-implantation genetic diagnosis (PGD) to test embryos for specific genetic conditions, and the Scientific and Clinical Advances Advisory Committee keeps abreast of clinical and scientific developments that the HFEA may need to consider. Because of the small number of HFEA members and the fact that they all maintain full-time jobs, the work of the sub-committees and the main committee is heavily supported by an Executive staff. Despite these various components of the institution, major decisions about the permissibility of new reproductive technologies or techniques are the responsibility of the core committee. As such, I focus largely on the decision-making of this group of appointed members, referring to it simply as ‘the HFEA’ for ease.

⁸⁵ Department of Health, “The Human Fertilisation and Embryology Act 1990: As Amended - an Illustrative Text,” November 9, 2007, http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_080206.pdf.

⁸⁶ The term ‘super-lay member’ was deployed self-referentially by a former member of the HFEA in interviews I conducted for a separate project in 2012.

The 2008 amendments to the HFE Act require the HFEA to maintain a statement of principles for its operation, but the principles that the Authority sets out are quite broad and procedural in nature:

1. We treat people and their information with sensitivity, respect and confidentiality,
2. We observe the highest standards of integrity and professionalism in putting into effect the law as it governs our sector,
3. We consult widely – listening to and learning from those with an interest in what we do,
4. We keep abreast of scientific and clinical advances, and
5. We exercise our functions consistently, proportionately, openly and fairly.⁸⁷

Following the third principle, the HFEA is well known for its consultations with the general public in formulating policy on controversial technologies and practices, like sex selection, financial compensation for egg and sperm donors, and genetic testing to select against certain diseases or conditions. It relies on a variety of methods for this public consultation, including polls, questionnaires, focus groups, and regional public forums. It also consults more extensively with patients, prospective parents, donor-conceived individuals, and providers of fertility treatment.

Because the HFEA “takes a pluralist approach to ethical questions,”⁸⁸ it does not formally adopt a specific ethical approach or more substantive principles than the ones listed above for its regulatory decision-making, although the HFE Act clearly dictates a number of substantive values. In 2011, the now-defunct Ethics and Law Advisory subcommittee⁸⁹ identified the following moral principles, drawn primarily from the HFE

⁸⁷ Human Fertilisation and Embryology Authority, “Our Role as Regulator,” last updated April 8, 2015, <http://www.hfea.gov.uk/135.html>.

⁸⁸ HFEA Ethics and Law Committee, “HFEA Principles” (London: HFEA, July 2003), http://www.hfea.gov.uk/docs/ELC_1_july03.pdf.

⁸⁹ The functions of the Ethics and Law Advisory Committee were passed on to the Ethics and Standards Committee in March 2013, which in turn was disbanded in January 2015 due to the current small size of the HFEA. Its functions will now be handled directly by the larger HFEA. See Human Fertilisation and Embryology Authority, “Ethics and Law Advisory Committee (ELAC),” last updated December 5, 2013,

Act and the Warnock Report, as those guiding the ethical, legal, and social advice it provided to the larger HFEA:

1. The **special status of the human embryo**, which limits research on embryos to the first 14 days after fertilisation and requires all such research to address a question relevant to a serious medical condition;
2. **Respect for persons**, encompassing informed consent, privacy, confidentiality, and access to certain information for donor-conceived children;
3. The **welfare of the child** born as a result of assisted reproduction;
4. **Reproductive autonomy**, although this is limited by the prohibition on non-medical sex selection, the prohibition on favouring genetically abnormal embryos, and considerations about the welfare of the child;
5. Furthering the **public good**, through public health measures, medical research, and public dialogue on fertility matters;
6. **Community, altruism and social trust**, which requires promoting practices that address the needs of those with medical or fertility problems; and
7. **Justice, equality and non-discrimination** in access to treatment.⁹⁰

Although not formally endorsed by the HFEA, this framework provides a good guide to the range of values that it has taken into account in its past policy reviews.

However, these general principles are frequently in tension. For example, deciding what kinds of embryo research to license – particularly whether the medical condition in question is ‘serious’ enough to warrant it – requires adjudicating between the special status of the human embryo and the benefits that may flow to the public or specific patient populations from this research.⁹¹ Similarly, considering the welfare of the child often conflicts with the principle of reproductive autonomy, as in the HFEA’s previous decisions to prohibit the selection of embryos based on sex or ‘abnormal’

<http://www.hfea.gov.uk/ELAC.html>, and HFEA Ethics and Standards Committee, “Minutes of the Proceedings of the HFEA Ethics and Standards Committee” January 14, 2015, http://www.hfea.gov.uk/docs/Ethics_and_Standards_Committee_Meeting_-_Signed_Minutes_-_Jan_15.pdf.

⁹⁰ HFEA Ethics and Law Committee, “Ethics Framework,” February 2011,

[http://www.hfea.gov.uk/docs/2011-06-15_-_ELAC_ethics_framework_for_website_\(2\).pdf](http://www.hfea.gov.uk/docs/2011-06-15_-_ELAC_ethics_framework_for_website_(2).pdf).

⁹¹ Human Fertilisation and Embryology Authority, “Decision Tree for Use by Licence Committee and Executive Licensing Panel Consideration of Applications to Grant or New a Licence,” n.d., http://www.hfea.gov.uk/docs/2014_004645__2014-03-06_-_decision_trees_collection_-_A3_size_pdf_-_final.pdf.

features like deafness and dwarfism.⁹² Public health efforts to reduce the multiple birth count by limiting the number of embryos transferred during IVF may also clash with the reproductive autonomy of those who prefer to have twins.⁹³

B. The practical challenge of moral reasoning in the arm's length context

As these cases suggest, in reaching decisions about when and under what circumstances citizens may access assisted reproductive technologies and techniques, the HFEA must make difficult moral judgments on behalf of the public. Even with a general set of substantive values endorsed by Parliament and handed down to the HFEA, individual policy questions require determining which of these principles apply in any given case and how to resolve tensions that arise between them. Situated at arm's length to ensure that these judgements are politically impartial and informed by the right kinds of expertise, members are expected to deliberate impartially and independently to a well-reasoned position. However, as a public body regulating an area as fundamental as reproduction, they are also expected to consult with and remain accountable to the public – the people who will be bound by these value judgments.

From the one side, the HFEA has been subject to criticism that it is not accountable enough in making such decisions, and its existence has been threatened on more than one occasion on this basis. In 2002, the House of Commons Select Committee on Science and Technology carried out an inquiry on developments in human genetics and embryology.⁹⁴ Questioning the former HFEA chair, Ruth Deech, soon after the

⁹² Julian Savulescu, "Deaf Lesbians, 'Designer Disability,' and the Future of Medicine," *BMJ* 325, no. 7367 (2002).

⁹³ Dominic Wilkinson et al., "Double Trouble: Should Double Embryo Transfer Be Banned?," *Theoretical Medicine and Bioethics* 36, no. 2 (2015).

⁹⁴ Science and Technology Select Committee, *Developments in Human Genetics and Embryology*, 10 July 2002, HC 791 2001-2002.

HFEA's decision to allow the creation of 'saviour siblings,'⁹⁵ one Committee member noted:

I wonder if Parliament really did intend, when it set up the HFEA, that very onerous and fundamental decisions that make basic changes in humanity should be decided by the HFEA under that legislation. We had a decision from the HFEA a few weeks ago to selectively, by genetic engineering, for the first time ever, give authority for the creation of a new human life for the specific purpose of benefiting an existing human life. That may be good or it may not be good. It certainly has enormous moral, ethical and social consequences... Should that decision have been taken by the HFEA or should it have been brought back to the democratically accountable Parliament to be discussed?⁹⁶

Deech countered by citing the importance of keeping such decisions at arm's length from Parliament, given the complex and rapidly developing nature of the field: such delegation "protects Members of Parliament from direct involvement in that sort of thing, which I may venture to say is right."⁹⁷ Furthermore, she contended that – because democratically accountable ministers appoint members of the HFEA after an open and transparent process – the HFEA is still a "democratic way" of making collective decisions.⁹⁸ The MP responded, "How can it be democratic if you are preventing the democrats, Members of Parliament who are elected to make difficult decisions on behalf of society as a whole, and protecting them from having to make such complex, fundamental decisions?"⁹⁹

This exchange crystallizes the legitimacy problem that arises in the arm's length governance of emerging biotechnologies: Deech defends the politically independent

⁹⁵ In the context of IVF, the term 'saviour sibling' refers to the practice of using pre-implantation tissue typing to allow parents to have a child who will be a tissue match for an older sibling with a life-threatening condition. See Human Fertilisation and Embryology Authority, "Pre-Implantation - Information on 'Saviour Siblings' and Testing for Genetic Conditions."

⁹⁶ Science and Technology Select Committee, *Developments in Human Genetics and Embryology*, 10 July 2002, HC 791 2001-2002, Ev 1.

⁹⁷ Ibid.

⁹⁸ Ibid.

⁹⁹ Ibid.

decision-making of the HFEA, but struggles to explain how these decisions of ‘enormous moral, ethical, and social consequence’ are simultaneously accountable to society. Meanwhile, the MP’s attachment to accountability as *electoral* accountability echoes the paradox position on the accountability-independence relationship: he implies that the only way for ‘complex, fundamental decisions’ to be made accountably in a democratic society is for them to be made by elected representatives. The HFEA, in other words, precisely because of its independence, is incapable of being accountable to the *demos* and thereby reaching a legitimate decision on fundamental, moral questions; indeed, this MP concluded his questions by suggesting that the HFEA may not be ‘fit for purpose.’¹⁰⁰

A decade later, the HFEA faced another existential challenge, as the Coalition government initially planned to abolish it under its public bodies reform programme and transfer its functions to a more general NDPB or agency, like the Care Quality Commission, or merge it with the Human Tissue Authority (HTA). Support for abolition came, on the one hand, from respected fertility experts like Lord Robert Winston¹⁰¹ and Professor Alison Murdoch,¹⁰² who argued that IVF should be regulated like any other medical treatment, and on the other hand, from Conservatives interested in reducing the size of the quango landscape and returning the HFEA’s functions to democratically accountable MPs. Echoing her predecessor Deech, Lisa Jardine lobbied heavily to save the body she chaired, arguing in the *BMJ*, “This is not an area where the average MP or civil servant should be asked to make decisions, let alone be in a position to make

¹⁰⁰ Ibid.

¹⁰¹ “Fertility Watchdog ‘Incompetent,’” *BBC*, December 10, 2004, sec. Health, <http://news.bbc.co.uk/1/hi/health/4084365.stm>.

¹⁰² Clare Dyer, “Independent Review Is Needed to Decide HFEA’s Fate,” *BMJ* 342 (2011).

decisions,”¹⁰³ and in the *Guardian*, “I believe these morally fraught issues must be held at arm’s length from government.”¹⁰⁴

Due to her efforts and a Department of Health consultation that received overwhelmingly negative responses to the planned abolition,¹⁰⁵ the Government agreed to hold off on abolishing the body until receiving the results of an independent review of the HFEA and the HTA. The review, while successfully advocating for the retention of each as separate NDPBs due to their distinct, important roles and generally excellent work, identified key concerns about the HFEA’s consultation methods:

There is also, however, a fairly widespread sense that the HFEA needs to do more to properly take into account stakeholder views and to be seen to do so¹⁰⁶...Many recognize HFEA’s expertise in handling large, set piece consultations...[b]ut this communication is seen as a predominantly one way process. There is not a general confidence that the organisation is really open to suggestions and ideas from outside, or that decision making processes are really influenced by stakeholders’ contributions.¹⁰⁷

This highlights the concern that although the HFEA’s consultations may be very publicly visible and collect a large amount of information on public views and values, it is not clear whether and how these results actually feed into decision-making. It is a criticism – like that of the MP in the Select Committee inquiry – of the HFEA’s failure to properly attend to public and stakeholder views. However, unlike the MP, the review seems

¹⁰³ Clare Dyer, “New Life for the HFEA?,” *BMJ* 342 (2011).

¹⁰⁴ Amelia Gentleman, “A Quiet Battle for Life: IVF Regulator Aims to Show Its Worth,” *The Guardian*, January 10, 2011, sec. Politics, <http://www.theguardian.com/politics/2011/jan/10/battle-life-ivf-regulator-hfea>.

¹⁰⁵ Department of Health, “Government Response to the Consultation on Proposals to Transfer Functions from the Human Fertilisation and Embryology Authority and the Human Tissue Authority,” January 25, 2013, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/212742/Consultation_HFEA_and_HTA_government_response.pdf.

¹⁰⁶ Justin McCracken, “Review of the Human Fertilisation and Embryology Authority and the Human Tissue Authority: An Independent Report to the Parliamentary Under Secretary of State for Public Health and the Minister for the Cabinet Office,” April 2013, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/216947/Justin_McCracken_report_of_review_of_HFEA_and_HTA.pdf, 15.

¹⁰⁷ *Ibid.*, 24.

hopeful that by revisiting its consultation strategy, the HFEA can adequately account for these views while simultaneously meeting the demands of “impartiality and independent action.”¹⁰⁸

The HFEA is no stranger, then, to accusations that it is not democratically accountable or that it fails to take public and stakeholder perspectives properly into account. However, from the other side, it has also been accused of pandering to public opinion and bending under political pressure at the expense of informed and robust moral reasoning. For instance, after the HFEA’s decision to ban social sex selection, discussed at greater length in Chapter IV, a significant number of UK and international bioethicists argued that the HFEA had merely endorsed the publicly popular policy option, throwing to the wayside its duty to consider the question rationally, impartially, and on the basis of the best scientific evidence.¹⁰⁹ Similarly, in a different report the Select Committee on Science and Technology criticized the HFEA’s consideration of sex selection in a separate inquiry, noting, “The issue requires greater analysis than has been afforded it by the HFEA”; although recognizing that the majority of the public oppose social sex selection, the Committee concluded, “Nevertheless, we do not see this as adequate grounds for prohibition...The onus should be on those who oppose [it] to show harm from its use.”¹¹⁰

Given the clarity of the accountability-independence tension in this instance, in which the HFEA is pulled in one direction to attend more to public input and in the other to insulate robust and rational deliberation from control by public opinion, how does the

¹⁰⁸ Ibid., 15.

¹⁰⁹ See John Harris, “Sex Selection and Regulated Hatred,” *Journal of Medical Ethics* 31, no. 5 (2005); Peter Herissone-Kelly, “The Prohibition of Sex Selection for Social Reasons in the United Kingdom: Public Opinion Trumps Reproductive Liberty?,” *Cambridge Quarterly of Healthcare Ethics* 15, no. 03 (2006); Edgar Dahl, “Sex Selection: Laissez Faire or Family Balancing?,” *Health Care Analysis: HCA: Journal of Health Philosophy and Policy* 13, no. 1 (2005).

¹¹⁰ Science and Technology Select Committee, *Human Reproductive Technologies and the Law*, 14 March 2005, HC 7-I 2004-2005, para 142.

HFEA practically approach these challenging moral judgments, and does this approach represent a legitimate way of making such public decisions? Because the HFEA formally endorses no particular framework for or method of reasoning, in contrast – as we will see below – to NICE, accounts of how it approaches this tension must be sought for in empirical studies of the HFEA’s decision-making process. Indeed, a large motivation behind this dissertation stems from the findings of a study of the HFEA that I undertook in 2012 on this question.¹¹¹ In that study, I used semi-structured interviews with members of the HFEA to construct an account of how the deliberative body considers different reasons and arguments in working towards regulatory decisions. I turn now to briefly explaining my findings and why they suggest the persistent absence of and need for an account of legitimate reasoning within this arm’s length context.

C. What method of reasoning for the HFEA?

In my 2012 qualitative study, I attempted to parse the HFEA’s *method* of reasoning: how certain factors and perspectives were recognized as ‘relevant’ within the HFEA’s deliberations, how arguments were evaluated in committee meetings, and how certain considerations ended up trumping others. At first, when I asked members how the HFEA uses its many ‘inputs’ in reaching a decision – e.g. public consultation, focus groups with patients and clinicians, and advice from its expert members from the fields of science, law, and philosophy – multiple members explained that the HFEA seeks to ‘balance’ views from these diverse inputs. For instance, one member argued:

The role of the Authority is to take everybody’s opinion and positions and thoughts into account and to balance them. I think that’s really crucial with everything that is done, to make everybody aware that we are looking at everybody’s perspective and that we’re there not on behalf of

¹¹¹ Kyle T. Edwards, “Methods of Legitimation: How Ethics Committees Decide Which Reasons Count in Public Policy Decision-Making,” *Social Science & Medicine* 113 (2014).

any particular group or in order to be swayed by any group but to have very sound reasons based on the whole range of thoughts as to the decision we make and why we make it.¹¹²

The idiom of ‘balancing’ is a recognizable part of everyday discourse, particularly in regards to ‘balancing various perspectives’ or ‘balancing different interests.’ However, while suggesting the attention that the HFEA dedicates to the broad range of viewpoints and beliefs represented in society, this account does not explain how the method of balancing yields a position on a given policy question.

When I asked members to explain what they meant by ‘balancing,’ they introduced the concept of ‘weight,’ the idea that certain perspectives or ideas are heavier or more powerful than others in the method of balancing. One member explained, “So simply because a particular group has formed which has a clear interest in the matter and campaigns on it and agitates and has material, I don’t think gives them any particular right or weight in terms of their views as opposed to members of the general public.”¹¹³ When prompted to explain which kinds of reasons or arguments carry more weight in this balancing process, members in turn identified considerations about who would be ‘helped’ or ‘harmed’¹¹⁴ by a decision or considerations voiced by those ‘affected’ by or ‘suffering’¹¹⁵ from an issue under consideration. In contrast to these welfare- or harm-based considerations, arguments about the ‘unnaturalness’ of certain reproductive technologies and practices, or about the proper relationship between parents and their children, were either seen to be less weighty or – particularly in the case of the former – not to be real or relevant reasons to be balanced amongst others.¹¹⁶

¹¹² Ibid., 37.

¹¹³ Ibid.

¹¹⁴ Ibid., 38.

¹¹⁵ Ibid., 37.

¹¹⁶ Ibid., 38.

This suggests an underlying pre-commitment to health and welfare concerns embedded within the HFEA's method of balancing, potentially at the expense of alternative social values – perhaps often seen as religious – like dignity, naturalness, or traditional family relationships. This is surprising given that 'public bioethics' bodies such as the HFEA were at least partially instituted in recognition of the shortcomings of the 'technical' or 'technocratic' model of scientific regulation; as explained by political scientist Alfred Moore, on this model scientists and clinicians, who either self-regulated or were entrusted via scientific advisory panels to set government policy on science and technology, largely based the limits of scientific inquiry or clinical practice on questions of risk and benefit, on "the presumption of ethical unity on the prioritization of health and welfare."¹¹⁷

As the Warnock Committee noted in proposing the creation of the HFEA, even if most can agree that there should be "*some principles* or other to govern the development and use of new techniques...[I]n our pluralistic society it is not to be expected that any one set of principles can be enunciated to be completely accepted by everyone."¹¹⁸ There will not be ethical unity in all cases of reproductive innovation, in other words, on the priority of health and welfare considerations (or others). As such, even those who tend to *agree* that these concerns should weigh quite heavily in the HFEA's decisions should recognize the need to publicly justify when and why these values are overriding; failure to make these normative judgements in a transparent manner hinders the HFEA's accountability to the public. Yet, currently, the apparent moral neutrality of the concept of 'balancing' belies the HFEA's substantive commitment to these principles.

¹¹⁷ Alfred Moore, "Public Bioethics and Public Engagement: The Politics of 'Proper Talk,'" *Public Understanding of Science* 19, no. 2 (2010): 201.

¹¹⁸ "Report of the Committee of Inquiry into Human Fertilisation and Embryology," 2.

We can draw two conclusions from this account of the way in which the HFEA reasons about whether to grant public access to novel biotechnologies. First, the HFEA needs a method of reasoning that does not mask, as the method of balancing does, the fundamental and often contested moral basis of the judgements it must make. Second, this method of reasoning, while providing a robust and independent approach to policy decisions, must also be suited to the public and collectively binding nature of such decisions: it is unlikely that any simple endorsement of primary principles – like health and welfare – will succeed in this respect. As such, the practical legitimacy challenge that I seek to address in relation to the HFEA is to locate a method of reasoning that allows the public body to make decisions that are, on the one hand, well informed and well reasoned, and on the other, publicly accountable: one that is capable of defending itself against past criticism that the HFEA is either too disconnected from the public perspective or too beholden to public opinion. I take up this challenge of articulating a legitimate method of reasoning for the HFEA in Chapter IV, based on the approach to public accountability that I set out in Chapter III.

2. National Institute for Health and Care Excellence

NICE was originally founded in 1999 as the National Institute for Clinical Excellence, at which point it was classified as a ‘special health authority,’ a type of arm’s length body sponsored by the Department of Health.¹¹⁹ It was tasked with deciding which drugs and treatments should be available on the NHS, through a technology appraisal process, as well as providing evidence-based guidance on treating a wide range of specific conditions. Most sources cite the main factor motivating NICE’s creation as

¹¹⁹ National Institute for Health and Care Excellence, “Who We Are,” *About Us*, <https://www.nice.org.uk/about/who-we-are>.

the increasing, negative attention throughout the 1990s to a perceived ‘postcode lottery’ in the UK healthcare system. This term is meant to capture the unfair variation in the types of treatments available to citizens based solely on the region in which they live.¹²⁰ In 2005, NICE merged with the Health Development Agency (HDA) and rebranded itself as the National Institute for Health and Clinical Excellence. The HDA had focused on gathering evidence on public health with the goal of reducing health inequalities and, as such, the merger heralded the increased involvement of NICE in guidance pertaining to public health and the active prevention of ill health.¹²¹ In 2013, it assumed its current title and was established as an executive NDPB as part of a set of changes brought about by the *Health and Social Care Act 2012*.¹²²

While most types of guidance that NICE produces are truly advisory in nature, the NHS is legally required to fund treatments approved through NICE’s technology appraisal programme if a patient’s doctor agrees that it is clinically appropriate.¹²³ Those technologies not recommended through this programme are rarely funded, as most technology appraisals are for new and expensive treatments rather than cheaper, existing ones, due to the role that appraisals play in allocating the limited resources of the NHS. Most emerging biotechnologies are flagged up for NICE by the National Institute for Health Research Horizon Scanning Centre at the University of Birmingham on a timeframe that affords NICE the opportunity to release a recommendation at the same time or soon after the technology comes to market.¹²⁴ Because the large majority of the

¹²⁰ Randeep Ramesh, “NHS Postcode Lottery Survey Reveals Wide UK Disparities,” *The Guardian*, December 9, 2011, sec. Society, <http://www.theguardian.com/society/2011/dec/09/nhs-lottery-survey-uk-disparities>.

¹²¹ National Institute for Health and Care Excellence, “Who We Are.”

¹²² *Health and Social Care Act 2012*, c. 7, <http://www.legislation.gov.uk/ukpga/2012/7/contents>.

¹²³ National Institute for Health and Care Excellence, “NICE Technology Appraisal Guidance,” <https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/nice-technology-appraisal-guidance>.

¹²⁴ National Institute for Health and Care Excellence, “Section 2: Selection of technologies,” in *Guide to the Processes of Technology Appraisal* (London: NICE, 2014), <http://www.nice.org.uk/article/pmg19/>.

UK relies upon the NHS as opposed to private healthcare, these binding technology appraisal recommendations can significantly extend or limit public access to new biotechnologies. In my treatment of NICE throughout this dissertation, I thus largely limit my focus to the technology appraisal process rather than other kinds of NICE guidance production, such as public health intervention guidance or best practice guidance for treating various conditions or diseases.

A. Decision-making principles and processes

In its charter, NICE cites seven ‘core principles’ that steer all work performed by the Institute. These principles require that NICE recommendations be:

- 1) supported by the best evidence of what works, in terms of both clinical and cost effectiveness,
- 2) created with expert input from lay members, representatives from clinical practice, public health, social care, and industry,
- 3) constructed with public involvement, by including two mandatory lay members on all committees,
- 4) produced by committees that are independent, transparent, and that conduct ‘genuine’ consultation,
- 5) subject to regular review,
- 6) based upon the social values of society, and
- 7) based upon internationally recognized methodologies for evaluating the clinical and cost effectiveness evidence discussed in the first principle.¹²⁵

In addition to these seven charter principles, the Institute’s *Social Value Judgements* document provides a set of both *procedural* principles for the design of NICE decision-making processes and *substantive* principles upon which to base individual decisions within these processes. To ensure that NICE’s substantive value judgements align with those of the public, this document draws upon reports of NICE’s Citizens Council, in addition to published bioethics literature, legislation on human rights, and expert

¹²⁵ National Institute for Health and Care Excellence, “NICE Charter,” April 2013, https://www.nice.org.uk/Media/Default/About/Who-we-are/NICE_Charter.pdf, 2-3.

input.¹²⁶ The Council is a group of 30 members of the public selected to roughly mirror the demographics of England and Wales (geographic location, occupation, sex, age, disability status, and ethnic minority status). As such, the focus is less on reaching a large portion of the public and more on using this smaller ‘mini-public’¹²⁷ as a type of citizens jury¹²⁸ to elicit and test the social values of the public.¹²⁹ As NICE’s primary method of public consultation, it meets for two days once per year to deliberate and ultimately report to NICE on a specific topic. Past reports have considered, for example, whether NICE should take the severity of a disease or the target population’s age into account in technology appraisals. While the eight substantive principles of *Social Value Judgements* are listed and discussed at greater length in Chapter V, the primary two (and those for which the Institute is most well-known) are that NICE should base its recommendations on the clinical effectiveness and cost effectiveness of interventions, using the unit of cost per quality adjusted life year gained (£/QALY) to compare novel technologies to alternative treatments. However, other substantive values, like avoiding discrimination and reducing health inequalities, are listed that temper these two.

Along with these substantive principles, *Social Value Judgements* endorses a set of procedural principles. Recognizing the long-standing tension between utilitarian and egalitarian approaches to the issue of distributive justice in healthcare, NICE avoids endorsing either one at the expense of the other. It instead opts for a procedural justice approach that requires decision-making committees to be accountable for their

¹²⁶ National Institute for Health and Care Excellence, *Social Value Judgements: Principles for the Development of NICE Guidance*, 2nd ed. (London: NICE, 2008): 6.

¹²⁷ Robert E. Goodin and John S. Dryzek, “Deliberative Impacts: The Macro-Political Uptake of Mini-Publics,” *Politics & Society* 34, no. 2 (2006).

¹²⁸ Jo Lenaghan, “Involving the Public in Rationing Decisions. The Experience of Citizens Juries,” *Health Policy* 49, no. 1–2 (1999).

¹²⁹ Michael Rawlins, “Background to NICE’s Citizens Council,” in *Patients, the Public, and Priorities in Healthcare*, ed. Michael Rawlins and Peter Littlejohns (Oxford: Radcliffe Publishing, 2009): 76–79.

reasonableness.¹³⁰ This ‘accountability for reasonableness’ approach, proposed by Norman Daniels and James Sabin, encompasses four procedural principles: publicity, relevance, challenge and revision, and regulation. ‘Publicity’ requires that the grounds for decisions are made public, ‘relevance’ requires that the grounds for decisions are ones that “fair-minded people” would agree are relevant, ‘challenge and revision’ requires the institutionalization of methods for contesting decisions once they are made, and ‘regulation’ requires that the decision-making process be monitored in order to ensure that it meets these conditions. I consider accountability for reasonableness in depth in Chapter III as one possible solution to the arm’s length legitimacy problem.

Guided by these substantive and procedural principles, NICE’s Technology Appraisal Committees must decide whether to recommend individual treatments for funding (see Figure 3). Evidence on the technology in question is submitted by ‘consultees,’ including the manufacturer, patient and carer interest groups, health professional societies, and clinical commissioning groups within the NHS. An evaluation report, which includes this evidence and an independent academic centre’s review of this evidence, is prepared prior to the meeting of the Appraisal Committee.¹³¹ The Committee considers the report, hears testimony from both clinical and patient experts in public meetings, and deliberates before reaching an initial recommendation. This is made available for comment, after which a final recommendation is made, accompanied by a thorough explanation required by the procedural principle of publicity. In line with the principle of challenge and revision, NICE maintains an appeals process in which these recommendations may be challenged. Barring any successful appeals, the appraisal is

¹³⁰ National Institute for Health and Care Excellence, *Social Value Judgements: Principles for the Development of NICE Guidance*, 9-10.

¹³¹ There are slightly different evidence review processes for single technology appraisal (STA) and multiple technology appraisal (MTA), but the distinctions are not necessary for my purposes. For more details see National Institute for Health and Care Excellence, “NICE Technology Appraisal Guidance.”

then published. If positively recommended, the NHS must make the technology available within three months.

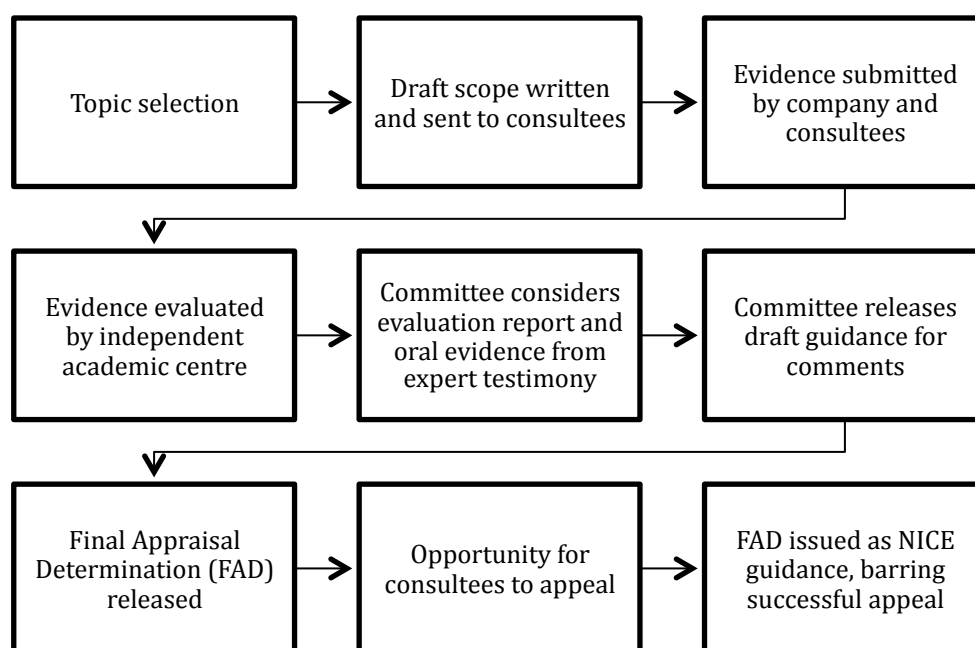


Figure 3. Technology appraisal process.

B. Matching 'greater autonomy' with 'increased accountability'

Having set out this basic account of NICE and its technology appraisal process, I turn now to examining how the legitimacy problem plays out in this setting. One of the most critical reasons for establishing NICE as an executive NDPB in the *Health and Social Care Act 2012* (HSCA) was to grant it a greater degree of independence in constructing guidance, in order to ensure that its decisions are not inappropriately influenced by the power of certain institutions or stakeholders.¹³² In the lead up to the passage of the HSCA, the main report on reforming the arm's length bodies associated

¹³² Department of Health, "Modernising Health and Care Public Bodies - The Health and Social Care Act 2012 Fact Sheet," June 15, 2012, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/138265/B6.-Factsheet-Streamlined-arms-length-bodies-250412.pdf.

with the NHS focused on safeguarding this independence. Titled *Liberating the NHS: report of the arm's-length bodies review*, this Department of Health report stated that reforms would “establish more autonomous NHS institutions, with greater freedoms, clear duties, transparency in their responsibilities to patients and their accountabilities, with power devolved to the front-line.”¹³³ To ensure that resource allocation decisions are made on the basis of the best available evidence with the interests of the entire population in mind, rather than on the basis of political pressure from powerful pharmaceutical manufacturers or interest groups, HSCA secured NICE’s institutional independence.

Equally, however, HSCA was passed in an effort to make the NHS and related arm’s length bodies like NICE “more responsive, efficient, and *accountable*” (emphasis added).¹³⁴ The Department of Health claims that the act ensures that “[g]reater autonomy for NHS bodies will be matched by increased accountability to patients, democratic legitimacy and a transparent system.”¹³⁵ Thus NICE is expected after these recent reforms to be at once more independent and more accountable. As set out in Chapter I, because the independence secured by the arm’s length location purposefully precludes classic methods of holding decision-makers to account, it is not immediately apparent what is required for NICE to be ‘accountable’ to patients or what it would mean for its operation to be ‘democratically’ legitimate. As in many such bodies, however, NICE has attempted to meet these demands through patient and public involvement, which diversify decision-making input beyond clinical and economic evidence and expertise.

¹³³ Department of Health, “Liberating the NHS: Report of the Arm’s-Length Body Review,” July 26, 2010, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/216278/dh_118053.pdf.

¹³⁴ Department of Health, “The Health and Social Care Act 2012 Overview Fact Sheet,” June 15, 2012, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/138257/A1.-Factsheet-Overview-240412.pdf.

¹³⁵ Department of Health, “Greater Accountability Locally and Nationally - The Health and Social Care Act 2012 Fact Sheet,” June 15, 2012, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/138264/B5.-Factsheet-greater-accountability-locally-and-nationally-250412.pdf.

NICE states, “By involving the very people for whom the guidance will be relevant, we put the needs and preferences of patients and the public at the heart of our work.”¹³⁶ Indeed, NICE maintains many formal pathways (some of which have been touched upon already) dedicated to ensuring that patients and the public are involved in technology appraisals (see Figure 4). The Citizens Council, through its impact on *Social Value Judgements*, and lay membership on Appraisal Committees and the Appeal Panel are the primary methods of involving the members of the ‘general public’ in technology appraisals. Organisations that represent specific groups of patients and carers relevant to the technology in question serve as consultees in the appraisal process, and may thereby:

- 1) comment on the scope of the technology appraisal, which importantly lists the outcomes that will be assessed to determine whether technologies are value for money,
- 2) submit evidence for the Appraisal Committee to consider,
- 3) nominate patient experts to attend the Committee meeting,
- 4) comment on the Committee’s draft guidance, and
- 5) appeal a Final Appraisal Determination.¹³⁷

Both patient and public involvement is supported by the Public Involvement Programme (PIP), a NICE team that helps create and sustain methods of involving patients, carers, and members of the public, as well as organizations that represent the interests of these groups.¹³⁸

¹³⁶ National Institute for Health and Care Excellence, “Putting Patients and the Public at the Heart of NICE’s Work,” *Public Involvement*, <https://www.nice.org.uk/about/nice-communities/public-involvement>.

¹³⁷ Lizzie Amis, “Patient Involvement in NICE Technology Appraisals,” in *Patients, the Public, and Priorities in Healthcare*, ed. Michael Rawlins and Peter Littlejohns (Oxford: Radcliffe Publishing, 2009): 30-35.

¹³⁸ National Institute for Health and Care Excellence, “Public Involvement Programme,” *Public Involvement*, <https://www.nice.org.uk/about/nice-communities/public-involvement/public-involvement-programme>.

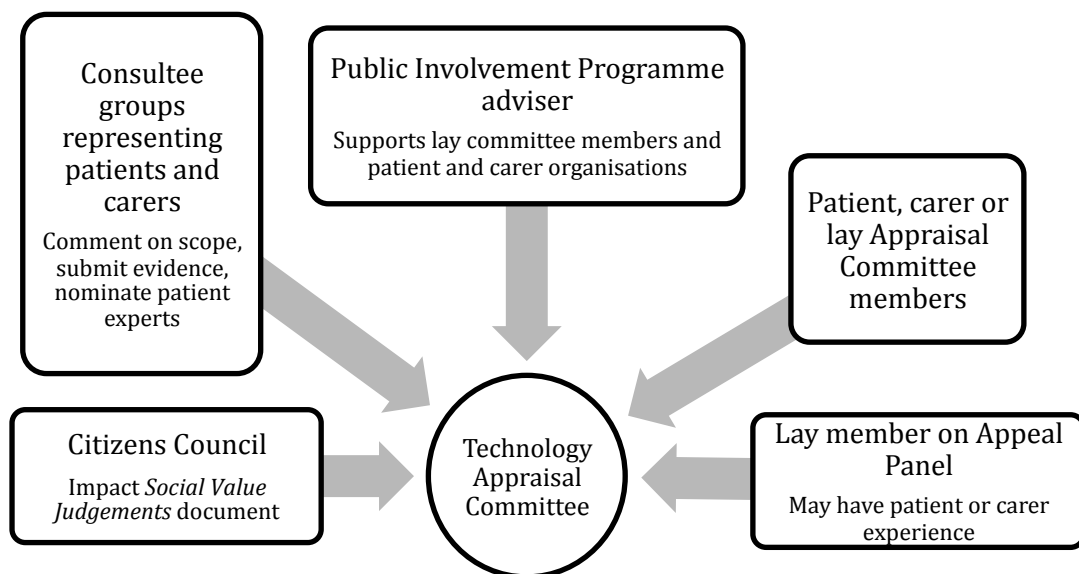


Figure 4. Patient and public involvement in technology appraisals.

Despite the existence of these clear, formal avenues for involvement, the *rationale* underpinning these avenues – particularly how they are meant to enhance accountability – is often ambiguous, particularly for the various methods of patient involvement. In turn, there is uncertainty within both NICE and the patient community about the role that patients should play in NICE’s technology appraisals. To illustrate the confusion surrounding this aspect of decision-making, I draw here on testimony before the House of Commons Select Committee on Health, the guidance that NICE provides to patients and the public, internal reviews by the Public Involvement Programme team, and independent qualitative studies of NICE.

C. What role for the patient voice in technology appraisals?

In preparing to appoint a new chair of NICE in 2013, the House of Commons Health Committee carried out an inquiry into the Institute’s work. The Committee’s

published report on this inquiry included a section dedicated to the ‘patient voice.’¹³⁹ Laura Weir, the Chair of Patients Involved in NICE, provided the main oral testimony on the current state of patient and public involvement. Patients Involved in NICE is a coalition of over 80 patient groups that is independent from but works closely with NICE. It maintains a formal memorandum of understanding with NICE and sees itself as a “critical friend and a respected and equal partner in developing and shaping aspects of NICE’s work.”¹⁴⁰ Indeed, the group holds its quarterly meetings in NICE’s offices and meets regularly with the Public Involvement Programme (PIP) team.

However, in spite of this close working relationship, Weir’s testimony before the Health Committee was quite critical of the institution’s PPI work, singling out the unclear role for patients:

With regard to patient group involvement in technology appraisals, there is no clear role for patient groups. We are invited to give oral evidence when NICE are approving a medicine and we are also invited to give a written submission before then. It is not clear whether NICE want us to literally comment on the assumptions going into their model or if they want a case study. Do they want to know how this drug has impacted on someone’s life? Do they want us to produce more qualitative or quantitative evidence? It is very unclear where we fit in the grand scheme of things...I wouldn’t say that patient groups are not included in the process; we are, but it is just that our role isn’t very clearly defined. Yet patient groups will spend considerable time and resources in terms of inputting into that process.¹⁴¹

Her testimony suggests that although formal avenues for patient involvement in NICE have been constructed, significant work remains for teasing out the details of why and how patient input affects technology appraisal decisions through those avenues.

¹³⁹ Health Select Committee, *National Institute for Clinical Excellence*, 16 January 2013, HC 782-I 2012-2013, para 49-51.

¹⁴⁰ “Memorandum of Understanding (MOU) Between National Institute for Health and Care Excellence (NICE) and Patients Involved in NICE (PIN),” November 2013, <https://www.nice.org.uk/media/default/About/NICE-Communities/Public-involvement/PIN/NICE-PIN-memorandum-of-understanding-November-2013.pdf>, 2.

¹⁴¹ Health Select Committee, *National Institute for Clinical Excellence*, 16 January 2013, HC 782-I 2012-2013, Ev 1.

Although demonstrably clear from these formal avenues of involvement that NICE wants to be properly accountable and responsive to patients, it remains unclear how patients fit into its approach to decision-making.

Weir is a stakeholder with a particular interest in strengthening the role that patients get to play in NICE decision-making, so it is necessary to look beyond her testimony in substantiating the confusion around the role of patients. The Chief Executive of NICE, Sir Andrew Dillon, voiced the belief in his testimony before the Health Committee that dissatisfaction from patient groups and other stakeholders about their involvement is at least “partly influenced sometimes by the nature of the outcome.”¹⁴² In other words, accounts like Weir’s may be biased due to the fact that, some of the time, NICE cannot recommend certain treatments that patients want on the basis of other considerations, like clinical and cost effectiveness. This is a crucial point to attend to in sorting out what it means for patient involvement to work ‘well’: given that technology appraisals are undertaken precisely because the NHS has limited resources, the measure of whether and how NICE is suitably accountable to patients requires a different metric than how often patients get what they want. This means that NICE needs a way to be responsive to patient input even – or particularly – when other considerations, like cost, are overriding.

Perhaps unsurprisingly, Dillon pushed back on Weir’s description of patient involvement in NICE, stating in his testimony:

I don’t agree that [the role of patients in the appraisal process] is not properly defined. It is very clearly set out in all of the programmes that NICE runs to produce guidance for the NHS, in the process documents and in the advice that we designed specifically for the patient organisations that we work with.¹⁴³

¹⁴² Ibid., Ev 24.

¹⁴³ Ibid.

This response mainly points back to the formal pathways for involvement, the existence of which is not disputed, rather than engaging with the question of how the patient voice should and does function within those routes. However, he also cites the advice that NICE provides specifically for patient groups. Does this advice provide a more specific explanation of the role of patient involvement?

The Public Involvement Programme team does indeed provide a large amount of material geared towards breaking down the complex processes and methods of technology appraisal for laypersons and informing patient groups about how they can participate at various stages.¹⁴⁴ These guides, however, are largely dedicated to interpreting more technical documents into plain English, rather than substantively addressing the types of information that patient groups should provide and how that information will be used. Partially on the basis of Weir's comments and the recommendations of the Health Committee, the team has since produced a set of 'hints and tips' documents that cover both the nomination of patient experts by patient groups and how to prepare to act as a patient expert. A review of these documents yields some additional specifics of what NICE expects from patient involvement:

1. Patient experts may be individuals with personal experiences of the condition or treatment in question or individuals that can act as a representative of such patients in virtue of "a thorough understanding of what matters to patients."¹⁴⁵ However, NICE discourages the

¹⁴⁴ See NICE Public Involvement Programme, "Scoping Technology Appraisals: A Factsheet for Patient and Carer Organisations," April 2014, <https://www.nice.org.uk/Media/Default/About/NICE-Communities/Public-involvement/Developing-NICE-guidance/Scoping-technology-appraisals-factsheet-patient-carer-organisations.pdf>; NICE Public Involvement Programme, "Developing Technology Appraisal Guidance: A Factsheet for Patient and Carer Groups," April 2014, <https://www.nice.org.uk/Media/Default/About/NICE-Communities/Public-involvement/Developing-NICE-guidance/Developing-technology-appraisals-factsheet-patient-carer-groups.pdf>; NICE Public Involvement Programme, "Overview of Technology Appraisal: A Factsheet for Patient and Carer Organisations," April 2014, <https://www.nice.org.uk/Media/Default/About/NICE-Communities/Public-involvement/Developing-NICE-guidance/Overview-of-technology-appraisals-patient-carer-groups.pdf>.

¹⁴⁵ NICE Public Involvement Programme, "Hints and Tips for Nominating Patient Experts for a Single Technology Appraisal (STA)," July 2014, <http://www.nice.org.uk/Media/Default/About/NICE-Communities/Public-involvement/Developing-NICE-guidance/sta-tips-nominating-patient-experts.pdf>, 3.

- nomination of a clinician as a patient expert, because the perspective of clinicians is already well represented in the appraisal process.
2. Although nominated by patient groups, patient experts act as individuals at appraisal meetings and not as representatives of the nominating group.
 3. NICE seeks patient expert nominees from different subgroups, in order to access the diversity of perspectives and experiences in the given patient/carer population rather than just the most common ones.
 4. The team notes that “evidence from Patient Experts tends to have the highest impact if it represents a balanced view,” where the expert can offer both benefits and drawbacks of the treatment in question.¹⁴⁶
 5. When responding to clinical issues in the appraisal process, patient experts are asked to consider whether their condition and the problems and benefits of the technology in question are accurately represented.
 6. When responding to cost-effectiveness issues in the appraisal process, patient experts are asked to consider whether the quality-of-life measures in clinical trials are reflective of what matters to patients.

Templates for patient groups’ evidence submissions request similar information: experiences of the condition and treatment, advantages and disadvantages of treatment, and the quality-of-life outcomes that matter to patients. They additionally ask whether certain subgroups may be expected to benefit more or less from receiving the treatment in question and whether there is any existing research on patient views of the condition or treatment.¹⁴⁷

While this provides practical advice for patients and patient groups, it falls short of explaining how it is that this type of information actually impacts the appraisal process. The Public Involvement Program (PIP) team itself, in a report on a survey of patient experts, notes, “[Patient experts] felt well supported by PIP in their preparation for committee, but they felt that they would like more help understanding their role at committee.”¹⁴⁸ Clarifying what one is being asked to do as a patient expert or patient group, in other words, is importantly distinct from clarifying *why* one is being asked to

¹⁴⁶ Ibid.

¹⁴⁷ “Patient/carer organisation consultee template for Single Technology Appraisal (STA)” from National Institute for Health and Care Excellence, “NICE Technology Appraisal Guidance.”

¹⁴⁸ Lizzie Amis and Heidi Livingstone, “Technology Appraisal Patient Expert Survey 2012 Report” (NICE Public Involvement Programme, April 2014), <http://www.nice.org.uk/media/default/About/NICE-Communities/Public-involvement/Public-involvement-programme/Patient-expert-TA-report-final-1.pdf>, 3.

do it. While NICE has admirably worked towards the former, the latter requires a different type of explanation, one that goes to the heart of *why* patients are involved rather than just *how* they are involved. Indeed, one of the key recommendations coming out of this report was for the PIP team to clarify the patient expert role in concert with the Technology Appraisal team, given that those surveyed, “[D]o not quite understand how they fit in given the clinical and technical nature of discussions and the length and weight given to these aspects during the meeting...They do not see clearly whether or how their contribution has informed the final outcome.”¹⁴⁹ This suggests that the general confusion may stem from questions about how the patient voice is and should be positioned relative to technical evidence about clinical and cost effectiveness.

To examine the relationship between these two inputs, it is useful to turn to two of the largest independent qualitative studies of the technology appraisal process. In 2001, Pauline Quennell conducted a project specifically on patient involvement in NICE technology appraisals, for which she carried out 29 interviews with patient group representatives, members of Appraisal Committees, NICE Board members, academics, and Department of Health staff.¹⁵⁰ And in 2004-2005, Timothy Milewa and Christine Barry carried out 33 interviews with Appraisal Committee members, NICE, NHS, and Department of Health staff, patient experts, and clinical and technical experts as part of a project to understand the role that patient and public concerns play in NICE’s decisions.¹⁵¹

¹⁴⁹ Ibid., 42.

¹⁵⁰ Pauline Quennell, “Getting Their Say, or Getting Their Way? - Has Participation Strengthened the Patient ‘Voice’ in the National Institute for Clinical Excellence?,” *Journal of Management in Medicine* 15, no. 3 (2001); Pauline Quennell, “Getting a Word in Edgeways? Patient Group Participation in the Appraisal Process of the National Institute for Clinical Excellence,” *Clinical Governance: An International Journal* 8, no. 1 (2003).

¹⁵¹ Timothy Milewa and Christine Barry, “Health Policy and the Politics of Evidence,” *Social Policy & Administration* 39, no. 5 (2005); Timothy Milewa, “Health Technology Adoption and the Politics of Governance in the UK,” *Social Science & Medicine* 63, no. 12 (2006); Timothy Milewa, “Representation

The interviews provide useful insight into how Appraisal Committee members view and interpret patient and patient group input. One Appraisal Committee member in Quennell's interviews argued quite bluntly that patient groups need to generate more robust evidence in order to impact decisions: "I happen to think that lay organisations are much more powerful if they can create the evidence base...if you want to influence health policy, you have to be able to demonstrate that the intervention's had some impact."¹⁵² This seems to ask patients groups to act more like researchers or pharmaceutical companies in collecting quantitative proof of impact.

Similarly, responding to what was, in 2001, a relatively novel NICE policy to allow patients themselves to testify before the Appraisal Committee, one staff member said:

I'm still fairly concerned about the presence of patients *per se* as opposed to individuals who are the advocates of those patient groups...the patients themselves feel this is their chance to express an opinion on this important process. From my point of view, I say that your representatives who deal with you day-to-day, who are advocates for your particular group, should be able to do that for you, and *would maintain a degree of professionalism as opposed to emotional advocacy* in the appraisal process (emphasis added).¹⁵³

Milewa and Barry, in the same vein, found that no less than 12 interviewees were frustrated with the emotional input received from patients, such as one Committee member who complained that some groups:

...just give you the entire textbook about the charity and the condition they deal with and why it is absolutely vital that you ignore the cost-effectiveness data (because, morally, that is what is expected of you).

and Legitimacy in Health Policy Formulation at a National Level: Perspectives from a Study of Health Technology Eligibility Procedures in the United Kingdom," *Health Policy* 85, no. 3 (2008).

¹⁵² Pauline Quennell, "Getting a Word in Edgeways? Patient Group Participation in the Appraisal Process of the National Institute for Clinical Excellence," 42.

¹⁵³ *Ibid.*, 42-43.

They might as well have torn up the paper and thrown it in the bin rather than post it to us.”¹⁵⁴

These accounts suggest that patients have simply misunderstood their role in the process: being accountable to patients, these members argue, does not and cannot require Appraisal Committees to accommodate emotional appeals or mere strong belief expressed in patient expert testimony or patient group submissions. Notably, while explicit about what patient involvement is *not*, none gives a positive account of the patient role.

Despite acknowledgement of these statements, patients have quite a different take on how their input impacts decision-making – if it does. Many patients think that their consultation submissions and patient expert testimonies rarely affect decisions: if the clinical and cost-effectiveness data are strong and clear, patient testimonies are “fairly irrelevant,” according to a patient expert in Milewa and Barry’s study.¹⁵⁵ Echoing this in the more recent 2012 Public Involvement Programme’s patient expert survey report, the team noted the “recurring theme that the patient role was seen as tokenistic.”¹⁵⁶ This concern was exemplified by one patient expert, who suggested, “In this context my contribution makes no difference because it lacks ‘objectivity’ against the RCT [randomised controlled trial] evidence, and patients, and all those except health economists, are effectively shut out of the discussion of the cost effectiveness of evidence.”¹⁵⁷ Even Committee members who were more receptive to patient group input than those cited above suggested that testimony from patient experts serves merely as an interpretive lens for the more determinative quantitative evidence: it helps turn data into “an extremely persuasive explanation” of the value of certain technologies, as opposed to

¹⁵⁴ Milewa and Barry, “Health Policy and the Politics of Evidence,” 503.

¹⁵⁵ *Ibid.*, 505.

¹⁵⁶ Amis and Livingstone, “Technology Appraisal Patient Expert Survey 2012 Report,” 27.

¹⁵⁷ *Ibid.*, 26.

being information that could impact or challenge the clinical and cost effectiveness data.¹⁵⁸

In the rarer cases in which patients do feel as if their input has affected the Committee's decision, the asserted method of this impact sits in stark contrast to the accounts given by Appraisal Committee members. One patient group representative in Quennell's study argued, "Although NICE tells us they don't want anecdotal evidence, at the same time we know that in certain cases it's the anecdotal evidence that has won it at the Appraisal Committee meetings."¹⁵⁹ Echoing this tacit understanding of what 'works' in Appraisal Committee meetings, a patient expert in Milewa's study noted:

I've been told that NICE are not swayed by emotive arguments...And I've kind of ignored that because what else have I got as a patient group? What else have we got except to appeal on the basis of, you know, 'We wouldn't put a dog or a cat through the kind of pain that people with rheumatoid arthritis have to endure...So, you know, I don't apologise for using that kind of language in a letter to [the Chief Executive of NICE] because, as I say, what have I got to lose and what else have I got in my armoury?'¹⁶⁰

This kind of emotive appeal suggests that patients seek to provide an account of the experience of a condition that cannot be captured in 'harder' clinical and cost effectiveness data. They also seem to think, despite the protest of Committee members to the contrary, that it 'works': that something in their personal anecdotes and emotional appeals can and has convinced Committee members to change their decisions.

It is difficult to say to what extent either of these accounts of patient involvement is reflective of reality. The recent boom in the statistical modelling of NICE appraisal decisions does little to clarify this question. One 2006 study showed that, while clinical

¹⁵⁸ Milewa and Barry, "Health Policy and the Politics of Evidence," 503.

¹⁵⁹ Pauline Quennell, "Getting a Word in Edgeways? Patient Group Participation in the Appraisal Process of the National Institute for Clinical Excellence," 42.

¹⁶⁰ Milewa, "Representation and Legitimacy in Health Policy Formulation at a National Level," 359.

and cost effectiveness are indeed the main determinants of outcomes, the mere presence of patient group submissions has a substantial effect on the odds of an intervention being approved for routine use.¹⁶¹ The causal basis for this effect is unclear however, and the authors caution that, due to the very small number of appraisals that do *not* include a patient group submission, the finding could be due to chance.¹⁶² When a similar study was undertaken in 2014, with a larger sample of decisions, the only independent factors that had a significant effect on decisions were cost effectiveness and the type of condition in question.¹⁶³ Pointing to the fact that NICE rejections are significantly less likely for cancer and musculoskeletal diseases, the authors tentatively suggest that patient group lobbying with the support of political and public opinion pressure may contribute to the significance of the type of condition. A contemporary study found no such bias for cancer treatments,¹⁶⁴ although the institution of the Cancer Drug Fund suggests the success of lobbying outside the appraisal process for politically salient diseases.^{165, 166}

In sum, while NICE is clearly committed to involving patients in its technology appraisal decisions, a lack of clarity about the rationale for this involvement has led to substantial confusion around the role of patients and patient interest groups. “We’re always being told how important we are...and how NICE value our input,” as one of Quennell’s patient interviewees put it, “Yet we’re never told *how* they value our input

¹⁶¹ Helen Angela Dakin, Nancy J. Devlin, and Isaac A. O. Odeyemi, “‘Yes’, ‘No’ or ‘Yes, But’? Multinomial Modelling of NICE Decision-Making,” *Health Policy* 77, no. 3 (2006).

¹⁶² *Ibid.*, 365-366.

¹⁶³ Helen Dakin et al., “The Influence of Cost-Effectiveness and Other Factors on NICE Decisions,” *Health Economics*, September 23, 2014.

¹⁶⁴ Karin H. Cerri et al., “Decision Making by NICE: Examining the Influences of Evidence, Process and Context,” *Health Economics, Policy, and Law* 9, no. 2 (2014).

¹⁶⁵ Martin Duerden, “From a Cancer Drug Fund to Value Based Pricing of Drugs,” *BMJ* 341 (2010); Martin Buxton et al., “Reforming the Cancer Drug Fund,” *The BMJ* 349 (2014).

and *why* they value our input.”¹⁶⁷ This is a very proximate challenge for the practical operation of technology appraisals: what sort of input should patients and patient groups provide and how should Appraisal Committees use this input? Yet it is also an instance of the more general legitimacy problem: how can NICE be accountable to those patients acutely affected by a specific appraisal decision, given that it must make independent decisions on the basis of the best available evidence and taking the interests of the greater population into account? Answering the very immediate question of how to involve patients thus requires sorting through the broader question of what it means to make a legitimate decision at arm’s length to restrict access to new technologies. After setting out a framework for such decision-making in Chapter III, I will show in Chapter V how it can resolve this patient involvement confusion.

3. Medicines and Healthcare Products Regulatory Agency

The Medicines and Healthcare Products Regulatory Agency (MHRA) is an executive agency of the Department of Health responsible for regulating all medicines and medical devices in the UK. The MHRA was founded in 2003 by merging the Medicines Control Agency and the Medical Devices Agency, both of which were formed in the 1980s in response to public outcry over high-profile medical disasters, such as the tragic case of thalidomide, that suggested the need for public regulation.¹⁶⁸ With the passage of HSCA, the MHRA merged with the National Institute for Biological Standards and Control (NIBSC) and also became the home of the newly founded

¹⁶⁷ Pauline Quennell, “Getting a Word in Edgeways? Patient Group Participation in the Appraisal Process of the National Institute for Clinical Excellence,” 42.

¹⁶⁸ Medicines and Healthcare Products Regulatory Agency, “Medicines and Medical Devices Regulation: What You Need to Know” (London: MHRA, April 2008), <http://www.mhra.gov.uk/home/groups/comms-ic/documents/websiteresources/con2031677.pdf>, 3.

Clinical Practice Research Datalink (CPRD).¹⁶⁹ The former is responsible for the production of an estimated 90 percent of world standards in biological medicines and assures drug quality in the UK through testing and research, while the latter facilitates the production and provision of anonymised healthcare records for research.¹⁷⁰ This new super-agency, officially created in April 2013, has assumed the ‘MHRA’ title with three separate ‘pillars’: MHRA Regulatory (the former MHRA), NIBSC, and CPRD.¹⁷¹

In addition to these larger and more formal mergers, the MHRA also absorbed a relatively large number of former arm’s length bodies as newly instituted ‘expert advisory groups’ during the Coalition government’s public bodies reform programme. The following were all previously separate NDPBs that were classified as ‘abolished’ but now operate within the framework of the MHRA: Advisory Board on the Registration of Homeopathic Products, Committee on the Safety of Devices, Herbal Medicines Advisory Committee, and the Independent Review Panel for the Classification of Borderline Products. Furthermore, there are a number of NDPBs that survived the recent cull that also sit within the MHRA, like the Commission on Human Medicines and the British Pharmacopoeia Commission.¹⁷²

As with NICE, where I focus specifically on the technology appraisal process due to its special role in regulating public access to emerging biotechnologies, I limit my attention to the original MHRA, rather than examining the newly incorporated NIBSC and CPRD. As this original division of the MHRA states in describing its relationship to

¹⁶⁹ Medicines and Healthcare Products Regulatory Agency, “What We Do,” *About Us*, <https://www.gov.uk/government/organisations/medicines-and-healthcare-products-regulatory-agency/about>.

¹⁷⁰ Winnie Agbonlahor, “Interview: Dr Ian Hudson, MHRA,” *Civil Service World*, May 20, 2014, <http://www.civilserviceworld.com/articles/interview/interview-dr-ian-hudson-mhra-0>.

¹⁷¹ *Ibid.*

¹⁷² Department of Health, “Establishing New National Bodies - The Health and Social Care Act 2012 Fact Sheet,” June 15, 2012, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/138272/C7.-Establishing-new-national-bodies-270412.pdf.

NICE, “MHRA says a product can be sold while NICE says it can be bought (by the NHS).”¹⁷³ It is responsible for deciding whether or not to license novel drugs and medical devices and thus whether citizens will be able to legally purchase or be prescribed a vast range of innovative technologies in the UK, which is why I have identified it for inclusion in this project. For ease, I continue to refer to this medicines and medical devices licensing division of the recently merged MHRA as simply ‘the MHRA.’

A. Decision-making principles and processes

The main goal of the MHRA is to protect and promote public health by ensuring that medicines and medical devices “work properly and are acceptably safe.”¹⁷⁴ Its primary test for these two principles of 1) safety and 2) effectiveness is whether the benefits outweigh the risks of the medicines and devices it oversees. In subjecting drugs and devices to this test, the MHRA stresses that only “robust and fact-based judgements” underlie its decisions.¹⁷⁵ As such, and in contrast to the HFEA and NICE, it explicitly denies the existence of any value judgements that its decisions may require. Emphasising transparency in decision-making, the MHRA affirms that “a reasonable person reviewing our decision should understand the rationale” behind it, that its decisions should be open and transparent wherever possible, and that its processes should be ones “in which the

¹⁷³ Medicines and Healthcare Products Regulatory Agency, “What Principles Inform the Decisions?,” last updated February 7, 2008, <http://webarchive.nationalarchives.gov.uk/20080804194642/mhra.gov.uk/printpreview/defaultsplashpp/con2032522?dynamiclistquery=null&dynamiclistsortby=null&dynamiclistsortorder=null&dynamiclisttitle=null&pagenumber=1&title=what%20principles%20inform%20the%20decisions%3F&resultcount=10>.

¹⁷⁴ Medicines and Healthcare Products Regulatory Agency, “Medicines and Medical Devices Regulation: What You Need to Know,” 2.

¹⁷⁵ Medicines and Healthcare Products Regulatory Agency, “What Questions Should Be Considered to Ensure That Reasonable Decisions Are Made?,” last updated February 7, 2008, <http://webarchive.nationalarchives.gov.uk/20080804194642/http://mhra.gov.uk/Howweregulate/Whatquestionsshouldbeconsideredtoensurethatreasonabledecisionsaremade/index.htm>.

public can have confidence.”¹⁷⁶ The MHRA alludes to the potential clash between the way in which scientific experts and lay members of the general public understand the balance between risks and benefits, although it is posed as a question the relevance of which is still undecided: “If the public’s views are in conflict with robust and comprehensive scientific evidence,” its main document on principles for decision-making asks, “should they carry any weight?”¹⁷⁷ This is listed as a question to “be considered to ensure that reasonable decisions are made,” but the institution has no defined approach to the possibility.

The MHRA has three groups that help produce its regulatory decisions: 1) its Executive staff, which carries out many of the day-to-day functions of the agency, 2) a number of advisory committees, notably the Commission on Human Medicines and the Committee on the Safety of Devices (currently in transition to the Devices Expert Advisory Committee), which provide expert advice on whether medicines and devices meet the agency’s two guiding questions of ‘Does it work?’ and “Is it acceptably safe?” and 3) the Agency Board, which supervises decision-making.¹⁷⁸ Before a company can sell its medical product in the UK, it must receive a license or marketing authorization for the product. This must be preceded by clinical trials, which require MHRA permission if carried out in the UK. Medicines and medical devices have different regulatory pathways: medicines are assessed directly by the MHRA, while devices are assessed by private sector ‘Notified Bodies’ that are appointed and audited by the MHRA. However, the Commission on Human Medicines and, soon, the Devices Expert Advisory Committee construct the demands of these regulatory pathways and consider

¹⁷⁶ Medicines and Healthcare Products Regulatory Agency, “Medicines and Medical Devices Regulation: What You Need to Know.” 2.

¹⁷⁷ Medicines and Healthcare Products Regulatory Agency, “What Questions Should Be Considered to Ensure That Reasonable Decisions Are Made?”

¹⁷⁸ Medicines and Healthcare Products Regulatory Agency, “Medicines and Medical Devices Regulation: What You Need to Know,” 4.

hard cases that arise therein, for medicines and devices respectively. In order to monitor safety and efficacy post-marketing, particularly of novel medicines and devices, the MHRA operates a Yellow Card Scheme that allows patients and carers to self-report side-effects. It also assigns ‘black triangle’ labels to new medicines and vaccines, or those being used for new indications, to highlight their novelty and the need to closely monitor them for unexpected side effects.¹⁷⁹

As a member of the European Union (EU), the UK has agreed with other Member States to pass the power to make certain decisions on medical products to the European level, adding a complex, international dimension to the MHRA’s work. It must attend to both UK law and European Parliament directives and regulations regarding medicines and medical devices, although the latter supersede Member States’ laws and regulatory instruments. This leads to some confusion as national legislation is formally updated to reflect new EU decisions, as when the UK Parliament replaced the *Medicines Act of 1971* and 200 related statutory instruments with the *Human Medicines Regulations 2012* in order to consolidate the law on medical licensing and bring it up to date with EU directives.¹⁸⁰

The EU is committed to fostering an internal market that encourages innovation and competitiveness while simultaneously ensuring high standards for the safety and effectiveness of medical products. Medicines may be authorized at either the European or national level, depending on the type of medicine in question. Those receiving marketing authorization at the European level may be sold in any Member State, and even those receiving authorization from national ‘competent authorities,’ like the MHRA, may be considered for recognition in other Member States. For medical devices,

¹⁷⁹ Ibid., 6.

¹⁸⁰ Olya Melnitchouk and Alison McAdams, “New Regulations Simplify 40 Years of Medicines Law,” *Health Service Journal*, November 29, 2012, <http://www.hsj.co.uk/resource-centre/best-practice/qipp-resources/new-regulations-simplify-40-years-of-medicines-law/5051815.article>.

Notified Bodies across the European Union approve all but the lowest risk devices, which are self-certified by the manufacturer. The MHRA oversees the Notified Bodies located in the UK and is responsible for pharamcovigilance in the UK after devices, as well as medicines, go on the market. Although much work has been dedicated to ensuring the harmonization of standards for medicines and devices across the EU, this means that the MHRA must rely in part on the authorization or certification of foreign or international bodies in protecting UK public health.

Because of this two-tiered system of UK and European-level legislation, both this section and Chapter VI will cite directives, regulations, and guidance from the European Commission, on the one hand, and UK legislation and regulations from the MHRA, on the other hand. The two sources together create a patchwork of legislative and regulatory instruments that the MHRA draws upon in fulfilling its role of protecting and promoting public health. Important for understanding the relationship between the MHRA and the European Commission, the UK plays a major role relative to other Member States in directing the development of medicines and medical devices legislation. As such, the work of the MHRA and the Commission are on the whole in sync and mutually reinforcing.

B. Challenges in the medical devices division: novel biotechnologies and borderline problems

High profile failures of medical device regulation to protect public health, such as the cases of PIP breast implants,¹⁸¹ metal-on-metal hip replacements¹⁸² and vaginal mesh

¹⁸¹ NHS Choices, “PIP Breast Implants - NHS Choices,” last updated September 7, 2014, <http://www.nhs.uk/Conditions/Breast-implants/Pages/PIP-introduction.aspx>.

¹⁸² Medicines and Healthcare Products Regulatory Agency, “Metal-on-Metal (MoM) Hip Replacements - Updated Advice with Patient Follow Ups,” *Medical Safety Alert*, June 25, 2012, <https://www.gov.uk/drug->

implants,¹⁸³ have prompted both the MHRA and the European Commission to review their regulatory approaches to medical devices. While serious incidents in their own right, these cases also brought to the fore the more general problem of how the increasingly diverse and complex spectrum of the medical devices market demands an overhaul of the regulatory process. In this section, I examine the current restructuring of the Devices Division within the MHRA and the European Commission's ongoing revision of its Medical Devices Directives to illustrate the key challenges that have been identified in planning for reform. In particular, I focus on how emerging biotechnologies, or novel uses of old ones, are creating 'borderline problems' for the MHRA. In such borderline cases, there is practical uncertainty about how the Agency should apply its traditional regulatory framework, threatening its ability to safeguard public health: it is not always clear in these cases what it means for a product to be 'acceptably safe' or to 'work properly.'

In 2013, the MHRA asked Professor Terence Stephenson, Chair of the Academy of Medical Royal Colleges, to form a review group to examine the MHRA's Devices Division. The resulting Stephenson report,¹⁸⁴ issued in 2014, found the Division to be sorely under-resourced relative to the Authority's medicines sector and lacking strong lines of communication to the clinical community, its counterpart institutions throughout the EU, and other UK health bodies like the NHS. The report was sweeping in its

device-alerts/medical-device-alert-metal-on-metal-mom-hip-replacements-updated-advice-with-patient-follow-ups.

¹⁸³ "What's the Issue with Mesh Implants?," *BBC News*, June 17, 2014, <http://www.bbc.co.uk/news/uk-scotland-27887766>.

¹⁸⁴ Professor Terence Stephenson, "Expert Clinical Advice - MHRA Medical Devices: Report of the Independent Review on MHRA Access to Clinical Advice and Engagement with the Clinical Community in Relation to Medical Devices," July 18, 2014, <http://webarchive.nationalarchives.gov.uk/20141205150130/http://www.mhra.gov.uk/home/groups/comms-po/documents/news/con402542.pdf>.

recommendations and described as ‘epoch-making’ by the Chairman of the MHRA.¹⁸⁵ In particular, the group advocated the creation of a new Devices Expert Advisory Committee to replace the Committee on the Safety of Devices, whose large membership and infrequent meetings hindered its ability to respond agilely to adverse incidents and innovation alike. The smaller Devices Expert Advisory Committee will help implement the changes of the report’s remaining recommendations, all of which were accepted by the MHRA.¹⁸⁶ The recommendations span four areas: 1) the organization of clinical input, resources, and leadership, 2) the collection and use of device incident data, 3) communications and partnerships with patients, the public, clinicians, and relevant UK and European institutions, and 4) future developments and emerging challenges.

Across the entire report, but particularly within the fourth theme’s focus on the “safe introduction of new and innovative technologies into clinical practice,”¹⁸⁷ the review group returns repeatedly to the ‘blurring’ of key boundaries and distinctions as one of the main challenges currently facing the MHRA. Although not delineated as such in the report, I identify three distinct boundaries of concern for the MHRA suggested by the Stephenson Committee’s review: 1) the line between the use of medical devices in the clinical context and in the home context, 2) the line between medical devices and medicines, and 3) the line between ‘medical’ and ‘consumer’ products more generally. The importance of these blurred lines, and the challenges they pose for regulation, is

¹⁸⁵ MHRA Committee on the Safety of Devices, “Minutes of the Committee on the Safety of Devices” March 20, 2014, http://www.gov.uk/government/uploads/system/uploads/attachment_data/file/401677/20-03-14_-_CSD_final_approved.pdf.

¹⁸⁶ Medicines and Healthcare Products Regulatory Agency, “MHRA Response to the Independent Review on Access to Clinical Advice and Engagement With the Clinical Community in Relation to Medical Devices,” July 18, 2014, <http://webarchive.nationalarchives.gov.uk/20141205150130/http://www.mhra.gov.uk/home/groups/comms-po/documents/news/con437436.pdf>.

¹⁸⁷ Stephenson, “Expert Clinical Advice - MHRA Medical Devices,” 43.

reinforced by similar attention garnered by the European Commission in its on-going revision of the Medical Devices Directives.¹⁸⁸

On the first boundary, between the use of devices in the clinical setting and the home setting, the Stephenson report notes:

The boundaries between devices designed for use in a medical setting and those for home consumers is blurring as increasingly sophisticated products are being bought over the counter or on the internet for self-management at home. More and more complex devices are being used by less skilled people.¹⁸⁹

While at-home management of health conditions using devices and software that measure and collect patient information is on the rise across the population, such as glucose monitors connected to smartphone apps for diabetic patients,¹⁹⁰ the shift from institutionalized or acute care to home care is nowhere more apparent than for elderly patients. As the UK population ages, unprecedented pressure is being placed upon care homes and hospitals, which will struggle to house and treat this large cohort. Many elderly patients prefer the independence and comfort of living at home, but, because of advances in medicine that have led to a longer average lifespan, more and more of them must manage chronic conditions in the later years of life. These concerns have been similarly highlighted during the European Commission's medical devices revision.¹⁹¹

¹⁸⁸ European Commission, "Revisions of Medical Device Directives," last updated June 10, 2014, http://ec.europa.eu/growth/sectors/medical-devices/regulatory-framework/revision/index_en.htm.

¹⁸⁹ Stephenson, "Expert Clinical Advice - MHRA Medical Devices," 3.

¹⁹⁰ Joseph Tran, Rosanna Tran, and John R. White, "Smartphone-Based Glucose Monitors and Applications in the Management of Diabetes: An Overview of 10 Salient 'Apps' and a Novel Smartphone-Connected Blood Glucose Monitor," *Clinical Diabetes* 30, no. 4 (2012).

¹⁹¹ European Commission, "Exploratory Process on the Future of the Medical Devices Sector: Potential Themes for Further Reflection at the European Level and Issues Identified by the Members" (Brussels, January 28, 2010), http://ec.europa.eu/health/medical-devices/files/exploratory_process/final_report_en.pdf, 7; European Commission, "Safe, Effective and Innovative Medical Devices and in Vitro Diagnostic Medical Devices for the Benefit of Patients, Consumers and Healthcare Professionals," Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, COM (2012) 540, (Brussels, September 26, 2012), <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52012DC0540&from=EN>, 3.

New innovations in telehealth may address these issues in tandem, by reducing the burden on traditional care pathways and enhancing patient autonomy by providing the option to live at home safely. Wearable devices, for instance, can measure and report vital signs to a doctor or carer, alerting them when measurements pass certain preset thresholds, and homes can be outfitted with devices that detect falls. Because of the benefits to both individual patients and population-level healthcare resources, the MHRA recognizes that it must encourage innovation in this area. However, this means that we can expect the blurring of the line between the clinical setting and the home to increase, not only as patients and lay carers (like family members) begin to use traditional medical devices in this new context, but also as new monitoring devices increasingly intertwine the home with the medical setting. The borderline problems that arise from this blurred line relate to how this shift to the home context affects the way that medical devices should be regulated: how should the concepts of risk, benefit, and effectiveness be understood here and which regulatory pathway should such borderline products be subject to?

This first blurred boundary, between the use of medical devices in home context and in the clinical context, lies *within* the category of medical devices: the central question it poses is not whether a product is or is not a medical device but whether and how the MHRA should treat devices differently depending on the context in which they are used. In contrast, the second boundary that is becoming increasingly fluid is that between medical devices and medicines. The Stephenson report highlights how “hybrid products are being developed which combine medicine with a delivery device (for instance medicated stents) and borders between medicines and devices are becoming

more blurred.”¹⁹² In theory, medicines and medical devices both fit into a broad category of medical products, the defining feature of which is to diagnose, treat, or prevent ill health. Whether any given product is classified as a medicine or a device *as such* should not affect the risk-benefit ratio that it is required to meet. However, in practice, determining whether medical devices work and are acceptably safe requires a different set of methods and expertise. This is one reason why medicines and medical devices are regulated under separate legislation and by separate groups at both the UK and European levels. For traditional medical products, like prescription pills or pacemakers, this division of pathways works well. However, hybrid products, in blurring the lines between medicines and devices, create confusion about which regulatory path a product should follow. In explaining one of the impetuses for its devices reform programme, the European Commission expresses the concern that unclear demarcations between the regulatory frameworks for medical devices and medicines, as well as other products like food and cosmetics, “may put patient safety at risk and lead to a fragmentation of the internal market.”¹⁹³

The final and arguably most challenging blurring referenced in the report is of the boundary between medical and non-medical devices:

There is increased blurring of expectations around ‘cosmetic’ devices and potential for a substantial increase in the number and nature of procedures. There is increased ‘consumerisation’ of the medical devices markets (Over The Counter, cosmetic, on-line, software apps).¹⁹⁴

[T]here is a blurred line between those types of software with a medical purpose and those with primarily ‘lifestyle’ applications.¹⁹⁵

¹⁹² Stephenson, “Expert Clinical Advice - MHRA Medical Devices,” 3.

¹⁹³ European Commission, “Executive Summary of the Impact Assessment on the Revision of the Regulatory Framework for Medical Devices,” Commission Staff Working Document, SWD (2012) 274, (Brussels, September 26, 2012), http://ec.europa.eu/growth/sectors/medical-devices_old/documents/revision/files/revision_docs/revision_ia_summary_en.pdf, 4.

¹⁹⁴ Stephenson, “Expert Clinical Advice - MHRA Medical Devices,” 19.

¹⁹⁵ *Ibid.*, 42.

While the second blurred line focused on the distinction between two types of medical products – medicines and devices – this line implicates the higher-level question of whether a product is ‘medical’ or ‘non-medical’/‘consumer’ in nature. As consumer products, including cosmetics, are regulated under separate legislation with different goals from medicines and medical devices legislation, the classification of a product as medical or non-medical is crucial in determining the type of scrutiny to which it will be subject. Medical products are identified by whether they make explicit or implicit medical claims, but uncertainty about what makes a claim ‘medical’ in nature is leading to more and harder borderline cases. Lifestyle smartphone apps meant to promote ‘healthy living’ and cognitive enhancement devices geared toward improving memory, for instance, both push the bounds of what we might traditionally think of as a ‘medical’ claim. Looking beyond the MHRA, just as the Commission emphasized the imprecise demarcation of medicines and medical devices, so too does it cite the regulatory gap of “implantable or other invasive products *without a medical purpose*” and the variable identification of direct-to-consumer genetic tests as medical devices across the European Union (emphasis added).¹⁹⁶

C. What regulatory boundaries for the MHRA?

Each of the three lines set out above represents a traditionally recognized boundary within the MHRA. The characteristics of new biotechnologies draw those traditional boundaries into question, ‘blurring’ them by raising problems with both the classification of individual products and the viability of the regulatory distinctions upon which the MHRA relies. Through the lens of the legitimacy problem, these borderline

¹⁹⁶ European Commission, “Executive Summary of the Impact Assessment on the Revision of the Regulatory Framework for Medical Devices,” 5.

products represent instances in which the MHRA is uncertain as to how to judge the safety and effectiveness of products, threatening its ability to make legitimate decisions on behalf of the public about which products should be permitted on or banned from the market. The focus on these borderline problems by both the MHRA's Stephenson review and the European Commission's review suggest the significance of this practical challenge to the traditional approach to regulation.

To address this challenge, I return in Chapter VI to these borderline products, focusing specifically on two of the blurred lines discussed here: that at the highest level between 'medical' and 'non-medical' products and that at the lowest level between different categories *within* the class of medical devices. The framework I endorse in the next chapter for making arm's length decisions about emerging biotechnologies provides a way to work through these borderline problems, but in so doing also suggests the need for a more general revision of the MHRA's fundamental regulatory boundaries if it is to continue to protect and promote public health in its licensing decisions.

Descriptively, this chapter provides the reader with a basic conception of the three arm's length bodies that serve as gatekeepers to emerging biotechnologies in the UK. More importantly, each section illustrates a current challenge to decision-making in each body and explains how this challenge is an instance of the legitimacy problem set out in Chapter I. The HFEA is often characterised as either being too removed from the views and values of the public in its decision-making on reproductive technologies or, conversely, merely following the tides of public opinion without adequate grounds for its decisions. It has not yet located a method of reasoning that would allow it to make independent, rational decisions in a publicly accountable fashion and thus decisively respond to these criticisms. NICE, on the other hand, endorses 'accountability for

reasonableness’ as its approach to decision-making on whether to fund expensive new treatments. Nevertheless, practical confusion about the role of patients and the patient voice in this process indicates residual uncertainty about how it may be simultaneously accountable to specific patient populations and impartial in the allocation of limited public resources. Finally, the MHRA has a longstanding and clearly defined approach to protecting public health by ensuring that the drugs and devices it licenses are 1) safe and 2) effective. However, a range of new devices do not fit easily into this regulatory approach, as evidenced by confusion around how to classify them and how to understand the concepts of safety and effectiveness in assessing them. These so-called borderline cases highlight the instability of some of the MHRA’s traditional regulatory boundaries, challenging the received wisdom of what it means to make a legitimate decision to license or restrict public access to drugs and devices.

These three accounts confirm that the tension between accountability and independence in the arm’s length context is not just a theoretical dilemma of mere academic interest: it causes genuine problems for real institutions that control decisions fundamental to citizens’ lives, and thus deserves the attention of bioethicists and policymakers seeking to improve the governance of biotechnology. In particular, it suggests how those already engaged in efforts to reform patient and public involvement in health governance or to articulate an account of ethics expertise in this area must consider the political features and legitimacy problem of the arm’s length context. Sorting out such practices within the individual institutions of the HFEA, NICE, and MHRA requires first attending to the roots of this problem that manifests in each.

The following chapter confronts the legitimacy problem as it has been set out in these first two chapters, turning to theories of public reason and deliberative democracy as potential ways of navigating it.

Chapter III. Navigating the Problem: Public Reason and Deliberative Democracy

As Chapter I illustrated, ‘quangos’ or arm’s length bodies are expected to fulfil a difficult task with little current guidance: they are expected to operate autonomously while still remaining accountable, to make decisions that are politically impartial while still acting in, and thus finding a way to conceive of, the ‘public interest’ in the consideration of scientifically complex and morally controversial issues. Chapter II then introduced the three arm’s length bodies that regulate public access to emerging biotechnologies, focusing on the way in which the legitimacy problem manifests itself in practical challenges within each institutional setting. Broadly, we want these regulators to be publicly accountable even as we insulate them from the most familiar method of being held to account in a democratic society, through the incentivizing and disciplining effect of elections for representatives.¹⁹⁷ Is this an impossible paradox, as one public official cited in Chapter I seemed to suggest, or can we find a robust way to conceptualize public accountability in the arm’s length context?

In fact, non-majoritarian institutions – those which “*by design* are not directly accountable to the voters or their elected representatives” – are not uncommon in democratic governments (emphasis added).¹⁹⁸ Especially well established in the United States, such institutions have more recently increased in prominence across Europe, both

¹⁹⁷ Robert J. Barro, “The Control of Politicians: An Economic Model,” *Public Choice* 14, no. 1 (1973); John Ferejohn, “Incumbent Performance and Electoral Control,” *Public Choice* 50, no. 1–3 (1986).

¹⁹⁸ Giandomenico Majone, “Europe’s ‘Democratic Deficit’: The Question of Standards,” *European Law Journal* 4, no. 1 (1998): 15.

in the structure of the European Union and within member states like the UK.¹⁹⁹ Just as the Cameron administration attempted to articulate with the Maude tests the conditions under which quangos are appropriate, political theorists have spent much time explaining when and why such arm's length institutions are both practicable and normatively legitimate within a democratic society. While a complete defence of delegation to non-representative bodies is beyond the scope of this dissertation, understanding its basic justification will be important in determining the circumstances under which arm's length governance is carried out in a way that is publicly accountable. There are two broad rationales for this delegation to independent institutions, both related to the way in which delegated governance makes decisions 'better,' albeit in quite different senses.

First, and most common, is a rationale for delegation based on the increased efficiency and effectiveness of the arm's length body. This rationale is largely analogous to the reason why we rely most of the time on representative rather than direct democracy. As citizens, we delegate the vast majority of political decisions to representatives because we reasonably expect that, based on their knowledge, experience, skill set, and increased time dedicated to considering political questions, government will function more efficiently.²⁰⁰ Delegation, then, is justified in part by the necessary time and the expertise in the day-to-day requirements of governing that representatives hold relative to the average citizen. In parallel, at one end of the spectrum of political decisions are those regulatory decisions, typically in areas like science or economics, that require increasingly complex technical expertise. Elected politicians will often delegate these decisions to arm's length bodies with the relevant expertise on the

¹⁹⁹ Giandomenico Majone, "The Rise of the Regulatory State in Europe," *West European Politics* 17, no. 3 (1994); Mark Thatcher and Alec Stone Sweet, "Theory and Practice of Delegation to Non-Majoritarian Institutions," *West European Politics* 25, no. 1 (2002).

²⁰⁰ Eric Maskin and Jean Tirole, "The Politician and the Judge: Accountability in Government," *American Economic Review* 94, no. 4 (2004): 1048-1049.

understanding that experts will more effectively and efficiently handle them. In cases like this, where delegation is based on technical expertise of some kind and focused on reaping efficiency gains, elected representatives in theory still set the broad policy goals that expert-driven arm's length bodies are expected to implement, on the understanding that the underlying values should reflect those of the public and are best determined by democratically accountable officials even if the details are set by technocrats. In other words, politicians are responsible for 'standard-setting,' determining the norms that will guide policy based on their constituents' preferences and values, while technical experts are responsible for 'standard-using,' or working on the basis of these handed-down standards to meet society's normative goals.²⁰¹ This first rationale clearly corresponds to the first of the Maude tests, that arm's length governance is appropriate for public bodies that perform a 'technical function.'

However, a second widely recognized reason for delegating decision-making to non-majoritarian institutions, rather than relying upon direct or representative democracy, is to protect against the 'tyranny of the majority.'²⁰² Courts with judges appointed for life, in contrast to political representatives elected for short terms, are the classic example of such institutions designed to protect those basic rights that should not depend upon the support of the current majority of the public. The recognition of the need to protect against the tyranny of the majority captures the more general understanding that what is popularly supported is not always justifiable: modern liberal democracies place certain constraints on the preferences of the majority to properly protect the liberties of individuals.

²⁰¹ This distinction between standard-setting and standard-using is drawn from Majone, "Europe's 'Democratic Deficit,'" 5-7.

²⁰² For the most famous account of the tyranny of the majority, see John Stuart Mill, *On Liberty*, ed. Steven M. Cahn (Rowman & Littlefield, 2005): 5.

In this second rationale for delegation, then, removing decisions from representative, majoritarian institutions is thought to make decisions *better* not in the sense of making them more efficient and effective, as was the rationale for delegation to technical experts, but in the sense of allowing due and proper reflection on the underpinning justifications for policy alternatives to ensure that they are fair and in line with the liberal commitments of modern democracies like the UK. Delegation allows this reflection to be unhindered by concerns of re-election and allows politicians to make credible commitments with longer time horizons than their short terms of office, especially when the benefits or justice of a controversial decision may take time to accrue or be recognized, respectively.²⁰³ This rationale underpins the second two Maude tests, of political impartiality and independence in establishing facts, and the few basic guidelines on ‘good governance’ for arm’s length bodies, overviewed in Chapter I, also seem based on this second rationale for delegation. Requirements that members be free of or at least declare conflicts of interest and that the appointment process for members must be open and publicized are based on the idea that delegated decision-making should not be distorted by any individual’s or group’s interests but rather guided by the long term good of society as a whole.

Importantly, unlike the case of delegation to technocrats for reason of efficiency and effectiveness, in which the normative framework for decisions might be largely set by elected politicians, this kind of delegation inherently requires a much greater cession of authority on the normative basis of policy. Here, in contrast to standard-*using* technical bodies, arm’s length bodies are standard-*setting*, dictating the norms or goals of policy in place of democratically elected representatives. This cession of authority

²⁰³ Thatcher and Sweet, “Theory and Practice of Delegation to Non-Majoritarian Institutions,” 4.

creates the perception of a ‘democratic deficit’²⁰⁴ and provokes outcry over the unaccountability of quangos that is so palpable in current politics. By removing important value judgments from control by citizens and their representatives, some worry that we are subverting the democratic process. However, once we recognize that democracy is properly conceived as more than a set of majoritarian decision-making processes, we need not understand this political delegation in the terms of a democratic deficit.

1. Deliberative versus aggregative conceptions of democracy

Philip Pettit, a central proponent of what has come to be called the ‘deliberative conception of democracy,’ argues for the depoliticization of representative democracy, by placing decisions in the hands of an arm’s length body, for issues in which “electoral interests are likely to militate against the *deliberative quality* of democratic decision-making, depriving *considerations of the common good* of the weight they are properly given (emphases added).”²⁰⁵ As he explains in justifying this apparently ‘undemocratic’ move, if democracy is understood as just an arrangement that allows the people by virtue of measures of popular control like referendums and a representative body to establish its own rules, then depoliticization would indeed be incompatible with democracy. It would place decision-making into the hands of unelected individuals or bodies.

This ‘aggregative conception’²⁰⁶ of democracy (also referred to as the ‘interest-based conception’²⁰⁷ or the ‘strategic conception’²⁰⁸), which defends broad use of the

²⁰⁴ Majone, “Europe’s ‘Democratic Deficit,’” 7.

²⁰⁵ Philip Pettit, “Depoliticizing Democracy,” *Ratio Juris* 17, no. 1 (2004): 53.

²⁰⁶ Joshua Cohen, “Procedure and Substance in Deliberative Democracy,” in *Deliberative Democracy: Essays on Reason and Politics*, ed. James Bohman and William Rehg (London: The MIT Press, 1997).

²⁰⁷ Samuel Freeman, “Deliberative Democracy: A Sympathetic Comment,” *Philosophy & Public Affairs* 29, no. 4 (2000).

majority principle, is both a well-known and popular interpretation of the demands of political equality. Describing the aggregative conception of democracy, which he identifies as “the reigning analytical conception,” Samuel Freeman explains:

Assuming that each person is to count as one, once a bare majority favours a proposal, then fairness demands that it carry the day according to the aggregative view. Otherwise, to require more than a bare majority is to assign greater weight to the preferences and interests of a minority...To conceive of democracy this way is to see it as serving a role similar to economic markets.²⁰⁹

In order to treat citizens as political equals on this account, then, what matters is the equal weight accorded to each citizen’s preferences in the determination of state action. On a purely aggregative conception, liberties would be unattainable unless the majority supported them: for the state to act otherwise would be ‘undemocratic,’ in the sense of failing to treat citizens as political equals. This means that liberties like freedom of speech and religion would not be inalienable: their continued protection would depend upon majority support from the current populace.

Freeman’s analogy of the aggregative conception with the market also deserves further attention. In his influential critique of such purely aggregative conceptions of democracy, Jon Elster uses social choice theory as an example of a market-based view of democracy, objecting to it because:

It embodies a confusion between the kind of behaviour that is appropriate in the market place and that which is appropriate in the forum...consumer sovereignty is acceptable because, and to the extent that, the consumer chooses between courses of action that differ only in the way they affect him. In political choice situations, however, the citizen is asked to express

²⁰⁸ David M. Estlund, “Who’s Afraid of Deliberative Democracy - On the Strategic/Deliberative Dichotomy in Recent Constitutional Jurisprudence,” *Texas Law Review* 71 (1992).

²⁰⁹ Freeman, “Deliberative Democracy,” 374.

his preferences over states that also differ in the way in which they affect other people.²¹⁰

Elster thus suggests a second feature of the aggregative conception of democracy: in addition to the way in which votes are counted, it has a particular view on the way in which citizens should vote. Or, more precisely, it refuses to prescribe a way in which citizens must vote, leaving them free, and arguably wise, to vote for alternatives that are in their self-interest.

In stark contrast to this conception of democracy is the deliberative conception, advocated for by Pettit, Freeman, and Elster. If, on the aggregative conception, legitimacy depends just upon the aggregation of citizens' preferences, what exactly distinguishes the deliberative view? The seemingly obvious answer, looking to its labelling, is the process of deliberation. Yet, as Freeman argues, even aggregative conceptions of democracy cannot possibly be understood as advocating that we register our votes without deliberating the consequences of the alternatives.²¹¹ Surely, to determine which option we prefer we must deliberate, at the very least internally, to determine how various policy alternatives affect our self-interest (or whatever end goal we choose to pursue).

It is not deliberation, then, that distinguishes the deliberative conception but rather “the *object* about which citizens deliberate and vote,” namely the common good (emphasis added).²¹² In other words, while aggregative conceptions are agnostic as to the ends that individuals must seek through voting,²¹³ deliberative conceptions require

²¹⁰ Jon Elster, “The Market and the Forum: Three Varieties of Political Theory,” in *Deliberative Democracy: Essays on Reason and Politics*, ed. James Bohman and William Rehg (London: The MIT Press, 1997): 10.

²¹¹ Freeman, “Deliberative Democracy,” 376-377.

²¹² *Ibid.*, 377.

²¹³ Note that there is nothing inherent to the aggregative conception of democracy that requires individuals to vote based on their own self-interest. They could very well choose to vote for alternatives on other grounds, including the common good.

citizens to base their voting behaviour on considerations of the common good. Legitimacy stems not from the mere fact of deliberation but from the types of ends that individuals pursue through the political system. It should be noted that deliberative conceptions need not be, and generally are not, conceived of as alternatives to voting. While consensus on a position that furthers the common good is the ideal outcome, most simply place certain conditions on the deliberation that must occur prior to voting. The significance of resorting to majoritarian procedures when consensus cannot be reached will be taken up later in the chapter.

Generally, as Pettit argues, direct or representative democratic institutions promote the identification of the common good and empowerment of public values – popular election of representatives, for instance, requires politicians to appeal to such common values in order to gain and retain office.²¹⁴ But when such democratic institutions frustrate this end goal, as in cases of moralism or instances in which the tyranny of the majority “carr[ies] the day against what commonly endorsed reasons would support,” such as those that sustain basic rights and liberties, they should be checked.²¹⁵ Arm’s length governance in a democracy is called for in precisely those instances in which electoral representation and control would tend to prevent the making of collective decisions on the basis of common, public values. And in such circumstances, setting decisions at arm’s length from elected politicians is far from antithetical to democracy or contributory to a democratic deficit. Rather, it is vital to its stability and integrity.

Furthermore, setting such decisions at arm’s length does not mean that we must accept the ‘paradox position’ and give up on accountability in equal measure to the

²¹⁴ Pettit, “Depoliticizing Democracy,” 59.

²¹⁵ *Ibid.*, 61.

independence that we grant to arm's length bodies in light of this deliberative goal. True, they may not be held accountable through traditional lines of democratic control, so called 'vertical accountability.'²¹⁶ But the rationale for their creation indicates the way in which their accountability may nevertheless be conceived. For technical arm's length bodies, established to bring efficiency and effectiveness to decision-making on preset normative goals, accountability may be measured through the ability to actually deliver these outputs. For those arm's length bodies established to bring political impartiality to decision-making, accountability may be conceptualized as the ability to accurately identify through deliberation the common good and public values upon which decisions should be based. In this way, we can conceive of public accountability beyond mere electoral accountability.

But what exactly do we mean when we refer to 'the common good' or 'the public interest,' and how are arm's length bodies to properly conceive of such ideas? These concepts seem elusive if not impossible given the diversity of the UK population. To answer these crucial questions, it is useful to turn to fleshed-out accounts of public reason and deliberative democracy, two sets of theories built upon the deliberative conception of democracy outlined here that seek to specify its requirements. Although generally directed towards modifying the practice of democratic governance broadly conceived, these theories provide promising approaches to the specific political problem of decision-making at arm's length, as each articulates how coercive political decisions may be publicly justified as normatively legitimate in a democratic society in the face of deep moral disagreement. Indeed, a variety of recent approaches developed to consider ethics in the context of health policy have drawn on public reason and deliberative

²¹⁶ Guillermo A. O'Donnell, "Horizontal Accountability in New Democracies," *Journal of Democracy* 9, no. 3 (1998).

democracy: Norman Daniels and James Sabin's accountability for reasonableness in resource allocation decisions,²¹⁷ James Fishkin and Robert Luskin's deliberative polling in eliciting the public's view on the NHS,²¹⁸ the myriad attempts at 'patient and public involvement' (PPI) at local, regional, and national levels,²¹⁹ and the Nuffield Council on Bioethics' 'public ethics' approach to regulating emerging biotechnologies.²²⁰

This chapter critically examines the potential of concepts drawn from deliberative democracy and public reason to guide arm's length decision-making on emerging biotechnologies, in a way that responds to the legitimacy problem by enabling bodies to make decisions that are both independent and publicly accountable. On the basis of this analysis, I argue for a process-oriented framework for public accountability in arm's length bodies based upon a condition of 'reciprocity' in public reasoning. However, my approach is highly conscious of and attentive to the persistent reasonable disagreement on the substantive basis of policies that emerge from this procedural framework. I begin by explaining the content of the two guiding theories themselves, focusing on seminal accounts by Amy Gutmann and Dennis Thompson, and John Rawls, as well as a number of alternate accounts and critiques levelled in response. I focus on the central similarities across the two sets of theories: the reasonable disagreement caused by the burdens of judgement and the resulting duty, captured by the 'criterion of reciprocity,' to rely on 'public' or 'mutually accessible' reasons in making political decisions.

²¹⁷ Norman Daniels and James E. Sabin, *Setting Limits Fairly: Can We Learn to Share Medical Resources?* (Oxford: Oxford University Press, 2002).

²¹⁸ Alison Park, Roger Jowell, and Suzi Mcpherson Contents, *The Future of the National Health Service: Results from a Deliberative Poll*, 1998.

²¹⁹ Julia Abelson et al., "Deliberations about Deliberative Methods: Issues in the Design and Evaluation of Public Participation Processes," *Social Science & Medicine* 57, no. 2 (2003); Susan Dodds and Colin Thomson, "Bioethics and Democracy: Competing Roles of National Bioethics Organisations," *Bioethics* 20, no. 6 (2006).

²²⁰ Nuffield Council on Bioethics, "Emerging Biotechnologies: Technology, Choice and the Public Good" (London: Nuffield Council on Bioethics, December 2012), <http://nuffieldbioethics.org/project/emerging-biotechnologies/>.

I argue that reciprocal reasoning ought to be a primary goal for arm's length bodies, which has important implications for public involvement and the way in which arm's length bodies relate to the public more generally. However, the extent and nature of disagreement in the very areas at issue in biomedicine – the start and end of personhood or the proper conception of distributive justice for healthcare, for instance – will often make the search for agreement on substantive 'reasons' unrealistic. In such cases, decisions will indeed require basing policy on reasons that are not fully 'public' or 'mutually justifiable' in nature. To make such decisions as legitimate as possible when public reasons are not forthcoming, we may set mutually acceptable procedural – rather than substantive – standards in place. These standards must properly recognize the gravity of making coercive political decisions on grounds that others may reasonably reject, by making the disputed substantive underpinnings transparent – not just the process itself – and creating avenues and fostering the means by which citizens may deliberatively contest these decisions.

2. Public reason and the criterion of reciprocity: giving up on the 'whole truth' in politics

Although public reason can trace its roots at least as far back as Kant's work in the 18th century, its revival and specification by Rawls in his 1993 *Political Liberalism* is the most well known account and has spawned a large literature of both support and criticism. Posing the central, motivating question behind this work, Rawls asks:

“[W]hen may citizens by their vote properly exercise their coercive political power over one another when fundamental questions are at

stake? Or in the light of what principles and ideals must we exercise that power if our doing so is to be justifiable to others as free and equal?”²²¹

Rawls’s answer to this question is set out in his ‘liberal principle of legitimacy,’ which states that the exercise of political power is proper only when based upon constitutional essentials that “all citizens as free and equal may reasonably be expected to endorse in light of principles and ideals acceptable to their common human reason.”²²² In recognizing others as free and equal, citizens must be willing to appeal to reasons that are justifiable to all other reasonable citizens who fall under the authority of the resulting decision, in other words ‘public reasons.’ The willingness to limit oneself to such reasons in public debate stems from what Rawls calls the ‘criterion of reciprocity,’ what I take to be the most important component of the idea of public reason and the most useful for understanding how the somewhat nebulous idea of ‘public reason’ may be specified.

Underlying the criterion of reciprocity, Rawls cites the ‘fact of reasonable pluralism,’ the fact that there exist a variety of *reasonable* moral, religious, and philosophical comprehensive doctrines that citizens endorse. As such, citizens will reasonably disagree about the nature of the good life on the basis of their conflicting but reasonable comprehensive doctrines. What explains this ‘fact’ of reasonable pluralism? This is important in distinguishing between reasonable and unreasonable comprehensive doctrines, as well as reasonable and unreasonable citizens. Rawls points to the ‘burdens of judgement’ as the reason why even reasonable citizens disagree.

²²¹ John Rawls, *Political Liberalism: Expanded Edition*, 2nd ed. (Columbia University Press, 2005): 217. Rawls limits the scope of public reason to “constitutional essentials and matters of basic justice” but others have since argued for a greatly expanded scope to include all coercive political decisions or indeed all instances in which we tell another how they ought to act. See Jonathan Quong, *Liberalism Without Perfection* (Oxford: Oxford University Press, 2010); Gerald Gaus, *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World* (Cambridge: Cambridge University Press, 2010).

²²² Rawls, *Political Liberalism*, 137.

These burdens of judgement are the factors that explain why reasonable citizens all seeking to reason rationally nevertheless frequently hold different and sometimes incompatible conceptions of justice and the good life. He lists six of these factors in *Political Liberalism*: 1) the complexity of empirical and scientific evidence relevant to an issue, 2) different weightings of relevant considerations even when reasonable citizens agree on this relevancy, 3) the difficult application of vague moral concepts to hard practical cases, 4) the diversity of experiences accumulated throughout our individual lives that affect how we understand practical questions, 5) the existence of strong normative considerations on either side of an argument that at least appear in practice to be incommensurable, and 6) the inherently limited range of even shared values that may be enshrined in any given social institution, which forces us to make difficult decisions about which to prioritize.²²³

Recognizing the burdens of judgement and the very different but reasonable beliefs that others in society may hold, reasonable citizens will not attempt to coerce others in structuring the rules of collective life on the basis of reasons drawn solely from their own comprehensive doctrines. Rather, citizens are defined as ‘reasonable’ when they desire to fairly cooperate with others in structuring public life and are willing to provide reasons for their proposals that others, who are similarly motivated, could accept. The criterion of reciprocity is this very willingness and motivation to seek laws that all could reasonably endorse, where ‘all’ is restricted to other citizens who also seek fair terms of cooperation on the basis of public reasons.

At the heart, then, of the idea of reasonableness is this criterion of reciprocity: the belief that we must not merely seek to show the truth or rationality of the principles we put forth for ordering society, but must rather justify them to others on grounds that they

²²³ Ibid., 56-57.

– so long as they are equally committed to this endeavour – could accept. Political liberalism, Rawls claims, is not concerned with the absolute truth – what one considers to be true or moral at a fundamental level – but rather “views this *insistence on the whole truth* in politics as incompatible with democratic citizenship and the idea of legitimate law (emphasis added).”²²⁴ In pluralistic societies, the criterion of reciprocity means that reasonable citizens will not appeal solely to their own comprehensive doctrines in making laws, as they recognize that treating others as free and equal requires giving reasons for the use of coercive political power that others, with conflicting comprehensive doctrines, can accept.

In contrast with the reasonable disagreement caused by the burdens of judgement, Rawls references those ‘sources of unreasonable disagreement’, such as prejudice, bias, self-interest (or group interest), blindness, and obstinacy.²²⁵ Unreasonableness is particularly conspicuous when individuals purposefully violate the criterion of reciprocity by refusing to seek fair terms of cooperation on grounds that others could reasonably accept. Unreasonable citizens, for instance, may come to the table unwilling to deviate from their understanding of the moral truth of an issue or may dissemble their true intentions by disavowing the political imposition of comprehensive doctrines only up until the point when they possess the power to impose their own. Unreasonableness, then, may refer either to the *motivations* of an individual, such as a person who attempts to legislate on the basis of his religious text in a pluralistic society, or to the *content* of a position, such as the prejudicial argument that only men should be allowed to vote. Jonathan Quong, a prominent scholar of Rawls and political liberalism, defines unreasonable citizens as rejecting at least one of the following central tenets: “1)

²²⁴ Ibid., 447.

²²⁵ Ibid., 58.

that political society should be a fair system of social cooperation for mutual benefit, 2) that citizens are free and equal, and 3) the fact of reasonable pluralism.”²²⁶ As such, any persons or doctrines that posit the greater moral worth of some persons (common examples include racism or sexism), or which refuse to recognize the reasonableness of certain views other than their own, or which impose beliefs drawn from essentially private experiences (such as faith or revelation²²⁷) are unreasonable.

As this brief overview of public reason may suggest, Rawls’s political liberalism and the idea of public reason invoke an interconnected linguistic web of ‘the reasonable’: reasonableness, reasonable citizens, reasonable pluralism, and reasonable disagreement. As Michael Saward has compellingly argued, and as the complexity of this web seems to suggest, Rawls is often ambiguous as to whether public reason is a *way* of reasoning or a *set* of reasons, a distinction between procedural and substantive accounts of public reasoning that is taken up in the next section.²²⁸ Suffice it to say for now that Rawls’ account of public reason is criticized by some deliberative democrats for entailing a ‘restricted agenda’ that is less a process of reasoning and more a principle imposing limits on reasons that may be voiced.²²⁹ On the one hand, the criterion of reciprocity seems to suggest that the idea of public reason is mostly about how we ought to reason – or the *process* for reasoning – when we have reasonable conflicting comprehensive doctrines. But to the extent that certain reasons are ruled out from the start – that public reason serves as a *substantive constraint* on reasons – the idea has faced a number of criticisms.

²²⁶ Quong, *Liberalism Without Perfection*, 315.

²²⁷ Erin Kelly and Lionel McPherson, “On Tolerating the Unreasonable,” *Journal of Political Philosophy* 9, no. 1 (2001): 44.

²²⁸ Michael Saward, “Rawls and Deliberative Democracy,” in *Democracy as Public Deliberation: New Perspectives*, ed. Maurizio Passerin D’Entreves (Manchester: Manchester University Press, 2002): 114.

²²⁹ Seyla Benhabib, “Toward a Deliberative Model of Democratic Legitimacy,” in *Democracy and Difference: Contesting the Boundaries of the Political* (Cambridge: Cambridge University Press, 1996): 75.

Consider public reason's uneasy relationship with religion. Many accounts of public reason, not just Rawls', distinguish between religious and secular reasons in politics and interpret the idea of public reason as excluding religious or religiously-motivated reasons from the set of justificatory public reasons. Robert Audi, for instance, argues that one should not support coercive legislation unless one has a 'secular' reason and motivation for its creation.²³⁰ Similarly, Bruce Ackerman suggests that one has a duty to justify the exercise of political power using 'neutral' reasons in a dialogue in which one cannot base the use of power on either the superiority of oneself or the superiority of one's conception of the good.²³¹ In general, this exclusion of religious reasons is based on the understanding that they are not 'shared' or 'publicly accessible' in a pluralistic society.

However, this account has received substantial criticism from those who respond that it is unfair to single out and exclude religious reasons in this way. Many non-religious reasons, they argue, are susceptible to many of the same challenges that religious reasons face:

What seems to us obvious is often a belief inherited from some religious authority, and it is accounted 'secular' only because the belief is widely shared. The point is not merely the historical one, that current secular beliefs were first adopted and enforced for 'religious' reasons, but that they actually make no sense, or have no reason, without particular metaphysical and epistemological assumptions that are, effectively, 'religious' (that is, they are taken on trust, are immune to empirical refutation, mingle ontological and emotional commitments, and are not universally endorsed by ordinarily rational beings).²³²

²³⁰ Robert Audi, *Religious Commitment and Secular Reason* (Cambridge: Cambridge University Press, 2000).

²³¹ Bruce A. Ackerman, *Social Justice in the Liberal State* (New Haven: Yale University Press, 1981).

²³² Stephen R. L. Clark, "Review of Religious Commitment and Secular Reason by Robert Audi," *Mind* 111, no. 443 (2002): 642–643.

This type of response suggests that the religious/secular binary is a heuristic of limited usefulness: many ‘secular’ reasons bear much in common with religious reasons to the extent that both may be based upon contested moral or metaphysical commitments. The secular reasoning of both a utilitarian and a non-consequentialist will ‘run out’ at the point in which their basic commitments are laid bare and in conflict – neither utilitarians, believing that actions should be assessed on the basis of their beneficial consequences, nor non-consequentialists, believing that certain duties take priority over the maximization of the good in guiding right action, have yet been able to decisively ‘prove’ the validity of their core belief. Yet this criticism only responds to the proposed stark division between religious and secular reasons. It does not attack the core premise of political liberalism, that the reasons used to justify the exercise of coercive political power must be shared in some sense by all reasonable people. It merely suggests that some secular reasons – those that are not shared – will, like religious reason, be excluded from this justificatory set.

A second set of challenges is instead directed towards this ‘shared’ criterion of public reason. Whether or not religious reasons are explicitly singled out, the demands of public reason mean that many citizens, often in practice religious citizens, will have to avoid drawing upon their most fundamental commitments and beliefs in advocating for or against various laws or policies. Some, including Michael Perry, argue that this represents an unfair restriction upon this class of citizens to the extent that beliefs about the human good ought to or even necessarily play an active role in public deliberation and political choice.²³³ Others, also concerned with the restrictive nature of public reason, suggest that it is not necessary to share in the *content* of another’s belief for it to

²³³ Michael J. Perry, *Love and Power: The Role of Religion and Morality in American Politics* (Oxford: Oxford University Press, 1993).

be admissible in public reason-regulated discourse; rather, it is the capacity to *empathize* with the fact or experience of that belief that matters.²³⁴ William Galston uses the example of a fundamentalist parent who objects to a component of a public school curriculum, arguing that even if others do not share and cannot relate to the religious basis for the objection, they should be able to empathize with the deep conflict the fundamentalist experiences between his beliefs about raising his child and the curriculum, and thus consider this conflict within deliberation rather than excluding it from the outset.²³⁵

A final kind of objection to the ‘shared’ criterion questions the supposed neutrality of the ‘reasonable.’ To be deemed public and capable of justifying coercive policies, reasons need only be shared by *reasonable* persons, a term which thus carries a great deal of power but, as suggested above, is difficult to pin down. Stanley Fish, for instance, suggests that words like reasonable, impartial, and disinterested all “clai[m] the high ground of neutrality while performing exclusionary work.”²³⁶ In other words, there is a risk that the reasonable may act as a kind of neutral-sounding mask that allows the speaker to smuggle in what is in fact a partial, substantive conception of the good, like the prioritization of health and welfare considerations seen in Chapter II’s discussion of the HFEA. As we will see, these critiques apply to conceptions of deliberative democracy and its corollary to reasonableness – mutual acceptability – as well, particularly in the notion that supposedly procedural accounts of legitimacy smuggle in substantive values that are not adequately justified.

²³⁴ William Galston, “Diversity, Toleration, and Deliberative Democracy: Religious Minorities and Public Schooling,” in *Deliberative Politics: Essays on Democracy and Disagreement*, ed. Stephen Macedo (Oxford: Oxford University Press, 1999).

²³⁵ *Ibid.*, 43.

²³⁶ Stanley Fish, “Mutual Respect as a Device of Exclusion,” in *Deliberative Politics: Essays on Democracy and Disagreement*, ed. Stephen Macedo (Oxford: Oxford University Press, 1999): 91.

3. Procedure and substance in public reasoning

Both public reason and deliberative democracy have arisen out of the so-called ‘deliberative turn’ in democratic political theory and are thus based on similar assumptions about what we owe other citizens in making political decisions and the failure of aggregative conceptions of democracy to meet these demands. Because of this common heritage, I stress in this chapter the similarities of theories of public reason and deliberative democracy, in particular, some articulation of a criterion of reciprocity as the central and shared component. Indeed, Rawls himself claimed in his later writings that the idea of public reason is a component of a deliberative democracy,²³⁷ and many identify him as a deliberative democrat.²³⁸ However, others deny that Rawlsian public reason can be understood as part of a deliberative democracy²³⁹ and focus on the distinction, rather than the similarities, between the two. One strand worth considering from this distinction is the debate over the relative superiority of procedural or substantive accounts of public reasoning.

Jürgen Habermas in his theory of deliberative democracy posits that legitimacy flows from democratic will-formation under conditions of free and inclusive deliberation, because under such conditions “no force except that of the better argument is exercised.”²⁴⁰ To the extent that actual deliberation mirrors the conditions of what he calls the ‘ideal speech situation,’ in which participants are treated as free and equal, fully informed, uncoerced, and are willing to have their minds changed by the force of the

²³⁷ Rawls, *Political Liberalism*, 448.

²³⁸ See Freeman, “Deliberative Democracy”; James Bohman and William Rehg, “Introduction,” in *Deliberative Democracy: Essays on Reason and Politics*, ed. James Bohman and William Rehg (London: MIT Press, 1997).

²³⁹ Simone Chambers, “Deliberative Democratic Theory,” *Annual Review of Political Science* 6, no. 1 (2003): 308; Saward, “Rawls and Deliberative Democracy,” 112.

²⁴⁰ Jürgen Habermas, *The Theory of Communicative Action: Reason and the Rationalization of Society*, trans. Thomas McCarthy (Boston: Beacon Press, 1985): 22.

better argument, the outcomes of deliberation will be rational and legitimate collective decisions. While participants may come to the decision-making table with interests and preferences, it is the deliberative process itself, as Seyla Benhabib explains, that generates their *considered* preferences.²⁴¹ Deliberators will modify their opinions or preferences on the basis of novel empirical evidence, the discovery of other relevant values or perspectives, and better arguments. It is only by properly structuring these ‘presuppositions of discourse’ that citizens may truly identify outcomes that are in the interest of all. Importantly, there must be no limits placed on the scope or content of deliberation: on Habermas’s account participants must “appeal to the best reasons *all things considered* from the universal point of view of ideal speech (emphasis added).”²⁴² This is in contrast to Rawls’ emphasis on appealing only to public reasons, rather than those drawn from moral, religious, or philosophical comprehensive doctrines, which Habermas would admit to deliberation.

Significantly, Habermas wants to distinguish his avowedly procedural account from what he defines as the substantive constraints of Rawls’ public reason. He sees himself as setting up the discursive rules of the game – setting up a process – and then ‘trusting’ actual deliberation amongst citizens to determine substantive outcomes, whereas he sees Rawls as significantly narrowing the scope of substantive outcomes by limiting the type of person that may deliberate ‘reasonably’ and the kinds of reasons that may carry weight. His theory, he argues, “[C]an leave more questions open because it entrusts more to the *process* of rational opinion and will formation...It leaves substantial questions that must be answered here and now to the more or less enlightened engagement of participants,” rather than to philosophers (like Rawls, the implication

²⁴¹ Benhabib, “Toward a Deliberative Model of Democratic Legitimacy,” 71.

²⁴² Freeman, “Deliberative Democracy,” 407.

goes).²⁴³ Under Rawls' quite stylized conception of 'reasonableness' and the 'reasonable person,' citizens "find themselves subject to principles and norms that have been anticipated in theory and have already become institutionalized beyond their control."²⁴⁴

As such, in a response to *Political Liberalism*, he notes that Rawls would do better if he "kept the procedural conception of practical reason free of substantive connotations by developing it in a strictly procedural manner."²⁴⁵ Rational opinion formation in the ideal speech situation, Habermas maintains, will weed out bad arguments and identify the good ones; so long as the 'unforced' force of the better argument is in action and the only force in action, other constraints are unnecessary. Furthermore, other substantive constraints will undermine the epistemic quality of results, by excluding potentially 'right' reasons, and the democratic nature of results, by limiting the extent to which actual deliberating citizens rather than isolated philosophical experts may determine political outcomes.

For his part, Rawls denies that Habermas can be as 'substance-free' as he suggests. As past failures of practical reasoning warn, citizens must have reference to substantive considerations and not rely purely on processes to deliver just results:

[O]ur considered judgments with their fixed points – such as the condemned institutions of slavery and serfdom, religious persecution, the subjection of the working classes, the oppression of women, and the unlimited accumulation of vast fortunes, together with the hideousness of cruelty and torture, and the evil of the pleasures of exercising domination – stand in the background as substantive checks showing the illusory character of any allegedly purely procedural ideas of legitimacy and political justice.²⁴⁶

²⁴³ Jürgen Habermas, "Reconciliation Through the Public Use of Reason: Remarks on John Rawls's *Political Liberalism*," *The Journal of Philosophy* 92, no. 3 (1995): 131.

²⁴⁴ *Ibid.*, 128.

²⁴⁵ *Ibid.*, 116.

²⁴⁶ John Rawls, "Political Liberalism: Reply to Habermas," *The Journal of Philosophy* 92, no. 3 (1995): 178.

In place of pure procedure, he argues, we must set limits on the types of reasons that participants appeal to, on the understanding that the ‘right’ considerations need to go in if we are to get the ‘right’ result. Directly responding to Habermas, Rawls argues:

The point is that no institutional procedure without such substantive guidelines for admissible reasons can cancel the maxim: garbage in, garbage out...[T]he mix of views and reasons in a vote in which citizens lack awareness of such guidelines may easily lead to injustice even though the outcome of the procedure is legitimate.²⁴⁷

By this, I take it he means that Habermas’s strictly procedural approach to reasoning to political decisions may fail to produce just results even if the process is apparently followed. Given that any actual deliberation will only be an *approximation* of ideal discourse, we ought to be suspicious of the claim of such outcomes to undeniable validity or justice, as practical shortcomings of the conditions of discourse may be expected.

The work of Joshua Cohen, a student of Rawls, suggests that the procedural and substantive elements of Rawlsian and Habermasian accounts of public reasoning may be somewhat reconciled. Properly structured public deliberation, he notes, does indeed lead to the identification of certain reasons and arguments as stronger than others, as Habermas suggests. Specifically, reasons and arguments that are responsive to the common good are those that are recognized as strong through the process. The nature of public deliberation is such that it “focuses debate on the common good,” since it is impossible to convince others in a truly free and inclusive discussion by arguing purely on the basis of self or sectional interest.²⁴⁸ If I want to persuade others of the validity of a claim, I must appeal to intersubjectively recognizable norms, and thus consider the types

²⁴⁷ Ibid, 178.

²⁴⁸ Joshua Cohen, “Deliberation and Democratic Legitimacy,” in *Deliberative Democracy: Essays on Reason and Politics*, ed. James Bohman and William Rehg (London: The MIT Press, 1997): 77.

of reasons that would convince others with interests and beliefs that diverge from my own. Public deliberation, in other words, acts as a sieve through which only arguments that are relevant to the consideration of the common good survive. And these reasons that survive the test of public deliberation will be aligned with Rawlsian ‘public reasons.’ Using this conceptualization of deliberation, he is able to suggest how if one accepts a deliberative democratic process that all may access and engage in as free and equal participants, “then one cannot accept as a reason within that process that some are worth less than others or that the interests of one group are to count for less than those of others. And these constraints on reasons will limit the substantive outcomes of the process.”²⁴⁹ This relationship between procedural and substantive principles of legitimacy is taken up again after the following, brief overview of deliberative democracy, and plays an integral role in my framework for public accountability.

4. Deliberative democracy: the persistence of deliberative disagreement in spite of reciprocity

I turn in this section to examining Amy Gutmann and Dennis Thompson’s account of deliberative democracy, which bears much in common with Rawls’ account of political liberalism and public reason but also departs from it at key points. Deliberative democracy as a modern area of political theory took off in the 1980s, largely as a response to the perceived deficiency of contemporary interest-based politics.²⁵⁰ Following Rawls, deliberative democrats of all varieties understand the basic democratic commitment of treating citizens as free and equal as requiring that we provide reasons or justifications to each other for the laws and institutions that structure

²⁴⁹ Cohen, “Procedure and Substance in Deliberative Democracy,” 415.

²⁵⁰ James Bohman, “Survey Article: The Coming of Age of Deliberative Democracy,” *Journal of Political Philosophy* 6, no. 4 (1998): 400.

collective life, particularly in the face of deep-seated disagreement about those laws and institutions. However, deliberative democracy pulls apart from the public reason tradition in tending to emphasize the importance of actual deliberation, often following Habermas and criticizing public reason liberals for what they take to be an individual project of hypothetical reflection on the types of permissible and impermissible reasons, rather than a shared project of engaging with other citizens and deliberating together in public.²⁵¹

Two of the motivations for Gutmann and Thompson's account are particularly helpful in contextualizing the main principles of their theory. First, Gutmann and Thompson question the practicability and desirability of the search for unanimity or consensus in deliberation. They base their principles of deliberative democracy on the belief that disagreement is "built into the conditions of collective life" and that politics, rather than aiming mainly to identify common ground, should take as its primary goal the search for the most justifiable way of making laws in the *absence* of agreement.²⁵² Of course, identifying consensus when it exists and resolving issues through the identification of previously unrecognized agreement is a valuable goal of the deliberative process, but their account stresses the vibrancy of pluralistic democracies and the value of working together in spite of the disagreement that thus arises.

Second, Gutmann and Thompson express frustration with the division between procedural and substantive theories of deliberative democracy and public reason discussed in the previous section. They present their theory as bridging this divide, although I do not find their explanation complete. The importance of this debate over

²⁵¹ See especially Amy Gutmann and Dennis Thompson, *Democracy and Disagreement* (Cambridge: The Belknap Press of Harvard University Press, 1996): 37; Amy Gutmann and Dennis Thomson, *Why Deliberative Democracy?* (Princeton: Princeton University Press, 2004): 4; Habermas, "Reconciliation Through the Public Use of Reason: Remarks on John Rawls's Political Liberalism."

²⁵² Gutmann and Thomson, *Why Deliberative Democracy?*, 27.

procedural and substantive principles is most apparent in cases where the outcome of a ‘fair procedure’ appears to violate what we consider to be a basic liberty, the concern raised by Rawls in his response to Habermas. At that point, we must either stick to the procedurally ‘just’ result or decide to uphold an apparently predetermined substantive commitment that has not been affirmed by citizens partaking in actual deliberation. Gutmann and Thompson usefully pose the issue that this presents to deliberative democrats like themselves: “How can deliberative democrats affirm substantive conclusions about politics and still support the value of actual deliberation, which may or may not produce those same conclusions?”²⁵³

They conclude that such conflicts are “similarly subject to deliberation and should accordingly be regarded as provisional. Neither substantive nor procedural principles have priority, even though citizens sometimes justifiably choose one over the other at some particular time.”²⁵⁴ While consistent with their commitment to subjecting the entirety of their theory to the process of actual and continual deliberation, this response appears to vacillate on the crucial question: under what circumstances should we refuse to accept deliberative decisions on the basis of substantive commitments or moral judgements, particularly as the stability of the deliberative system depends upon the ‘losers’ accepting – at least temporarily – laws that they consider to be based upon false or less than the most reasonable positions? Deliberative democracy, like public reason, requires that we give up the search for ‘the whole truth’ in the structuring of collective life, instead paying heed to what others can reasonably accept – except, apparently, when it comes to certain basic but unspecified rights. As these rights appear

²⁵³ Ibid., 121.

²⁵⁴ Ibid., 123.

prior to the deliberative democratic system, it remains unclear how it can determine and justify these external rights as Gutmann and Thompson would like it to.

These elements of dissatisfaction – with the search for consensus and the procedural/substantive divide – set the stage for Gutmann and Thompson’s own theory. Their account of deliberative democracy has three principles that regulate the process of politics and three principles that regulate the content of policies, highlighting their belief in the necessity of both procedural and substantive principles. The three procedural principles, which I focus on here, are 1) reciprocity, 2) publicity, and 3) accountability. Just as the criterion of reciprocity is the central component of public reason, so too is the principle of reciprocity the central component of deliberative democracy. It entails deliberative democracy’s ‘reason-giving requirement,’ which demands that citizens give each other reasons that are *mutually acceptable* for the political decisions they make.²⁵⁵ In addition to this constraint on the types of reasons that can justify political decisions, reciprocity requires that the empirical claims citizens make be subject to reliable and mutually acceptable standards of inquiry. It seems, then, that deliberative democracy begins the same way as public reason, with the criterion of reciprocity and the search for reasons that others can accept.

Importantly, particularly in light of the frequent demand for ‘impartial’ decision-making in arm’s length bodies, Gutmann and Thompson maintain that reciprocity is distinct from and preferable to impartiality. While reciprocity demands that our reasons be mutually acceptable, a principle of impartiality insists that whether others accept a position is irrelevant so long as it is true or correct from an impersonal perspective.²⁵⁶ Proponents of impartiality would thus be willing to coerce others on the basis of non-

²⁵⁵ Ibid., 3.

²⁵⁶ Gutmann and Thompson, *Democracy and Disagreement*, 59.

mutually acceptable reasons so long as they believed they had reached this impersonal perspective. In this way, impartiality is yet another ‘comprehensive morality,’ their analogue to Rawls’s comprehensive doctrine, and as such cannot be the basis of decision-making in a deliberative democracy.

Gutmann and Thompson depart from public reason by stressing that deliberative disagreement will often persist to the point at which decisions must be made and votes must be tallied. ‘Deliberative disagreement,’ as opposed to more generic forms of disagreement, is distinctive in that it characterizes the state of affairs that exist when citizens on each side of an argument are equally committed to finding fair terms of cooperation and have both given reasons for their position that do not appear to violate the principle of reciprocity’s requirement of mutually acceptable reasons. In other words, the claim that deliberative disagreement persists is a claim that the criterion of reciprocity and the dependence on public reasons are not always enough to yield decisions. The authors attribute the persistence of this deliberative disagreement to four features: scarcity of resources, limited generosity, incompatible values and incomplete understanding.²⁵⁷ These largely mirror Rawls’ burdens of judgement, which he suggested primarily to explain the existence of conflicting but reasonable comprehensive doctrines, but did not think would lead to an absence of public reasons or the inability to apply public reason in the way that Gutmann and Thompson suggest. I will take up this question of the limits of the criterion of reciprocity in the next section.

The second procedural principle of publicity asks that deliberative democracy’s mutually acceptable reasons be given in public during actual deliberation between citizens. This principle of publicity contrasts with the internal deliberation or private reflection of citizens using Rawlsian public reason. Rawls only requires that the grounds

²⁵⁷ Ibid., 21-26.

for decisions *could* be accepted by reasonable citizens; it could thus be determined privately by a philosopher engaging in an exercise of pure reason rather than in a public deliberation.²⁵⁸ Basing political decisions upon such private deliberation would preempt a variety of the goals of deliberative democracy, particularly the potential of deliberation to transform the beliefs and positions of citizens and to clarify the nature of persistent moral disagreement when such transformation ceases.

Gutmann and Thompson's third and final procedural principle is that of accountability, which relates both to who gives the mutually acceptable reasons required by the principle of reciprocity and to whom these reasons must be addressed. As outlined in Chapter II, arm's length bodies like the HFEA and NICE struggle to both come to well-reasoned decisions and, as suggested by the frequently unclear purpose of 'public consultations' and 'patient involvement,' properly take into account the views of the public to whom they are accountable. Gutmann and Thompson consider these problems of accountability in the terms of challenges posed from either end of the spectrum: deliberative accountability is often criticized for being either too populist, when representatives give too much weight to public opinion and merely reflect the preferences of constituents uncritically, or too elitist, when a particular class of deliberators come to dominate moral reasoning and political decision-making.²⁵⁹ The authors attempt to assuage these concerns by claiming that deliberative democracy "searches for modes of representation," rather than a single form of accountability, and does not endorse any particular way of being accountable in all situations. However, they do insist that decisions that implicate basic liberties require representatives to override majority opinion if they are convinced that it would violate these basic liberties,

²⁵⁸ Ibid., 37.

²⁵⁹ Ibid., 132-144.

endorsing in such cases deliberative reasoning over representation. Deliberation itself should determine how much deference to popular opinion is appropriate in circumstances that do not implicate basic liberties.²⁶⁰

5. Can public reason yield definitive answers for arm's length bodies?

A key problem with public reason suggested by Gutmann and Thompson's account is that it seems to prove 'incomplete' at times. Public reason may be incomplete in one of two ways: it may be inconclusive or it may be indeterminate.²⁶¹ Public reason is inconclusive when it "fails to generate convergence among reasonable people on a single political outcome."²⁶² This happens when participants, operating on the basis of reciprocity, argue for two conflicting policy positions that are nevertheless both based on values that appear to be mutually justifiable. In such cases, the use of public reason seems to produce two or more reasonable positions and its application cannot conclude which of these positions to choose. While deliberators may be able to argue that theirs' is the most reasonable position, multiple reasonable options exist. The classic example of such inconclusiveness is the application of public reason to abortion. Rawls argues:

[W]e consider the question [of a right to abortion] in terms of these three important political values: the due respect for human life, the ordered reproduction of political society over time, including the family in some form, and finally the equality of women as equal citizens... Now I believe any reasonable balance of these three values will give a woman a duly qualified right to decide whether or not to end her pregnancy during the first trimester. The reason for this is that at this early stage of pregnancy the political value of equality of women is overriding, and this right is required to give it substance and force.²⁶³

²⁶⁰ Ibid., 139-142.

²⁶¹ Gerald F. Gaus, *Justificatory Liberalism: An Essay on Epistemology and Political Theory* (New York: Oxford University Press, 1996); Micah Schwartzman, "The Completeness of Public Reason," *Politics, Philosophy & Economics* 3, no. 2 (2004).

²⁶² Schwartzman, "The Completeness of Public Reason," 192.

²⁶³ Rawls, *Political Liberalism*, 243.

Critics, however, argue that here Rawls clearly ignores the question of whether the foetus is a person. Citizens who believe that the foetus is a person could argue against a right to abortion on a different balance of these very same political values, that the value of respect for human life outweighs the value of the equality of women: “Some who agree with Rawls about the values at stake will claim that a foetal right to life is required to give substance and force to *that* value,” Phillip Quinn points out, just as Rawls thinks that a right to an abortion is required to give the political value of equality of women the proper force.²⁶⁴ Like Quinn, while I want to agree with Rawls’ conclusion, it does not seem that public reason alone can get us to it.

As a slightly different take on the incompleteness of public reason in the case of abortion, Robert George suggests that Rawls “seems plainly, if silently, to import into the analysis of the question a range of undefended beliefs of precisely the sort that ‘political liberalism’ is supposed to exclude,” namely by “smuggling in” the controversial moral and metaphysical claim that the foetus is not a person.²⁶⁵ He notes:

It is true that arguments in favour of a fetal right to life cannot avoid appealing to certain metaphysical and moral propositions that are currently in dispute among reasonable people. But the same is true of arguments for a right to terminate pregnancy and deliberatively bring about fetal death.²⁶⁶

To the extent that neither side of the argument can avoid appealing to such contested claims, he and others, like Kent Greenawalt²⁶⁷ and Michael Perry,²⁶⁸ think that controversial moral claims drawn from comprehensive doctrines are *required* in this

²⁶⁴ Philip L. Quinn, “Political Liberalisms and Their Exclusions of the Religious,” *Proceedings and Addresses of the American Philosophical Association* 69, no. 2 (1995): 43.

²⁶⁵ Robert P. George, “Public Reason and Political Conflict: Abortion and Homosexuality,” *The Yale Law Journal* 106, no. 8 (1997): 2488.

²⁶⁶ Robert P. George, “Law, Democracy, and Moral Disagreement,” *Harvard Law Review* 110, no. 7 (1997): 1392.

²⁶⁷ Kent Greenawalt, *Private Consciences and Public Reasons* (New York: Oxford University Press, 1995).

²⁶⁸ Perry, *Love and Power*.

instance to reach a position on the right to abortion, that there is simply no political position on abortion that does not violate the requirements of public reason. In other words, rather than generating multiple positions supported by reasons that all reasonable persons could accept, the attempt to meet the demands of reciprocity produces *no* public reasons capable of resolving the issue, only reasons based in contested comprehensive doctrines. This kind of objection suggests that public reason is *indeterminate* rather than merely inconclusive.

Gutmann and Thompson's explanation of deliberative disagreement is also a claim about the incompleteness of public reason, although it blurs the line on whether it is indeterminate or inconclusive. In the US debate on the legalization of abortion, they suggest, "Both pro-life and pro-choice advocates argue from fundamentally different but plausible premises to conflicting public policies. Both make *generalizable claims* that are also *recognizably reciprocal* in their moral and empirical content."²⁶⁹ In other words, at least some citizens on each side meet the demands of the criterion of reciprocity – they seek fair terms of cooperation, believe in the freedom and equality of all people, and base their arguments on the types of claims that others should be willing to accept. Although in many cases the criterion of reciprocity may generate a solution, in such cases Gutmann and Thompson admit that the principle of reciprocity by itself is incapable of locating a mutually acceptable resolution.

The worry about the application of public reason within arm's length bodies regulating emerging biotechnologies is that, even if public reason is only incomplete in a relatively narrow set of issues, questions about the permissibility and rationing of biotechnologies will often land precisely in that narrow set. This seems to be one of the key assumptions driving both the Nuffield Council on Bioethics' public ethics approach

²⁶⁹ Gutmann and Thompson, *Democracy and Disagreement*, 74.

to the class of emerging biotechnologies and Norman Daniels' and James Sabin's accountability for reasonableness approach to rationing healthcare resources. I here highlight the assumptions in each account about the particular features or characteristics of emerging biotechnologies that we may expect to fuel the incompleteness and ultimate insufficiency of public reason. I save their proposed solutions to this problem for consideration in the following section.

In 2012, the Nuffield Council on Bioethics released a report titled *Emerging biotechnologies: technology, choice and the public good* that seeks "to examine common social, ethical and legal issues raised by emerging biotechnologies" towards the end of "mak[ing] recommendations on research, policy, governance and public engagement."²⁷⁰ The Council justifies this broad consideration by noting that, despite the diversity of research programmes, forms of knowledge, practices, products, and applications that make up the large class of emerging biotechnologies, they warrant a common approach due to the problems associated with 'emergence' and with their shared focus on the 'biological' or living systems.

Emerging biotechnologies, the Council argues, share three distinctive characteristics posing a 'threefold' challenge that "make[s] it difficult to arrive at a universal rational basis for commitment to particular biotechnologies": the *uncertainty* about the range or likelihood of outcomes, the *ambiguity* of meaning and value attached to practices even if the expected outcomes are predictable, and their *transformative potential*, or the way in which certain biotechnologies may affect shared ways of life for all, not just for the 'users.'²⁷¹ Ambiguity is perhaps a poor term for the second characteristic, and indeed the Council clarifies, "By 'ambiguity' we mean a lack of

²⁷⁰ Nuffield Council on Bioethics, Terms of Reference, "Emerging Biotechnologies: Technology, Choice and the Public Good."

²⁷¹ Nuffield Council on Bioethics, Chapter 3: The Threefold Challenge of Emerging Biotechnologies, "Emerging Biotechnologies: Technology, Choice and the Public Good."

agreement about the implications, meanings or relative importance of a given range of possible outcomes...[which] reveals the association of different and possibly incompatible meanings and values.”²⁷² These features, particularly this second one of the incompatibility of meanings and values or different ordering of such values, point to why even sincere and prolonged public deliberation may not lead to a “universal rational basis” for a decision, a claim again about the limits of public reasoning in this context. Indeed, just as Gutmann and Thompson stress the persistence of deliberative disagreement in pluralistic societies, the Council assumes, “In a society that accommodates plural values, it is no longer plausible to suggest that the end of ethical reflection is a single, rational answer to the question ‘what is the good life?’”²⁷³

In particular, the Council stresses the way in which emerging biotechnologies, such as many advances in the area of reproductive technology that the HFEA must regulate, challenge categories and distinctions upon which people base their moral judgements, such as organism/artefact, human/non-human, born/made, and natural/artificial.²⁷⁴ Here we find reference to the way in which characteristics of biotechnologies may severely limit the scope for the use of public reason. Specifically, in cases where the very question at stake is who or what is a moral subject entitled to some level of moral consideration, cases that Greenawalt helpfully refers to as ‘borderline status cases,’ there may be no public reasons upon which to appeal.²⁷⁵ As in the case of abortion, many debates about pre-implantation genetic diagnosis, stem cell research, and the future of human germline genome modification seem to depend on questions about

²⁷² Ibid., 40.

²⁷³ Ibid., 62.

²⁷⁴ Ibid., 47.

²⁷⁵ Kent Greenawalt, *Religious Convictions and Political Choice* (New York: Oxford University Press, 1991).

the status of the embryo that – at least currently – we seem unable to address without appeal to nonpublic reasons.

Questions about which emerging biotechnologies, in particular new drugs but also new devices, should be covered by a public healthcare system like NICE also point to the limits of public reasoning, although usually in a different way than reproductive technologies. For healthcare allocation, novel biotechnologies prove especially difficult cases for allocation decisions that *already* suggest the incompleteness of public reason, largely because of their great expense and the lack of a ‘track record’ for their effectiveness beyond initial and often flawed trials. Indeed, although Daniels’ and Sabin’s accountability for reasonableness approach is directed towards the problem of healthcare resource allocation more generally, they suggest that the “major force driving up health care costs worldwide,” and thus the most immediate motivation for an account of how we can set limits fairly, is “the emergence of new technologies, including new procedures, drugs, devices, and diagnostic techniques.”²⁷⁶ Foregrounding their approach, Daniels and Sabin note:

Unfortunately, no democratic society we are aware of has achieved consensus on...distributive principles for health care. Even people who want to cooperate in steering a society’s health system will hold diverse moral and religious views. Their individual value systems will lead them to disagree morally about what constitutes a fair allocation of resources to meet competing health care needs.²⁷⁷

In other words, even if a group of deliberators are all committed to finding fair terms of cooperation on the basis of reciprocal reasoning, they will often continue to disagree about the proper substantive principles for decision-making. In particular, the application of public reason to such priority-setting decisions seems to be *incomplete*: even if

²⁷⁶ Daniels and Sabin, *Setting Limits Fairly: Can We Learn to Share Medical Resources?*, 1.

²⁷⁷ *Ibid.*, 2.

reasonable people can agree on mutually justifiable considerations – limiting themselves to evidence, reasons, and principles that they all can agree are *relevant*, in the terms of Daniels’ and Sabin’s accountability for reasonableness, there is still the problematic task of ordering conflicting values within this set of relevant considerations and clarifying how vague considerations should be interpreted. There is still “*reasonable disagreement* about what guidance they give for real-world decision-making (emphasis added).”²⁷⁸ This account of the problems of resource allocation suggests that public reasoning in this context is likely to be inconclusive rather than indeterminate. Daniels and Sabin seem fairly confident that people will be able to agree on relevant reasons (their analogue to public reasons), in contrast to the example of abortion and reproductive technologies where people may not be able to agree on relevant reasons in the first place.

So, for instance, reasonable people may agree that the relevant considerations in deciding whether a new drug should be covered by the NHS are its clinical effectiveness, its cost, and broader considerations of equality, but the relative importance of each of these considerations must still be settled. Firstly, people may disagree about what each consideration means: is clinical effectiveness the years of life gained compared with using the standard of care or years of life gained adjusted for their quality and, if the latter, what factors affect this quality? Arguably, the relevancy condition could be reapplied to each of these three broad considerations to specify them in line with public reason. But, and secondly, at some point the deliberators must decide whether a given drug is a proper trade-off between its cost and its effectiveness. And it seems that the search for public reasons, or identifying relevant reasons, will not be very fruitful in setting this line, both because of different weights individuals will attach to each and the

²⁷⁸ Ibid., 4.

arbitrariness of the decisions near the margins, even when people relatively agree on the ordering or appropriate weights of considerations.

Past attempts to institute a ‘simple solution’ to the problem of priority setting in healthcare on the basis of substantive rules or goals to be attained have failed, as Soren Holm explains, because of both a lack of consensus on such principles and disagreement over how abstract principles should be specified and weighed in practice.²⁷⁹ For example, a ‘first wave’ of *principle*-based priority setting in Norway, Sweden, Denmark, and the Netherlands beginning in the 1980s, in which independent commissions attempted to set out substantive principles to guide decisions, like the severity of disease, cost-effectiveness, equity, and solidarity, failed due to lack of agreement about which principles were relevant, how principles were to be applied, and what to do when the principles clashed. Even when all are committed to appealing to mutually acceptable reasons, it seems that disagreement persists. And if all we have are those inconclusive substantive principles (or reasons), we may feel that we are left with no recourse in justifying the ultimate decision we must make in such cases.

6. Where can we go when public reason ‘runs out’?

Micah Schwartzman claims that indeterminacy is the more damning type of public reason incompleteness, because it means policies must be based on claims that others can reasonably reject.²⁸⁰ Returning to the example above, either allowing or banning abortion implicates fundamentally contested metaphysical beliefs. Cases of reproductive technologies that invoke similar claims about the status of the embryo may also result in public reason indeterminacy. Conversely, inconclusiveness merely means

²⁷⁹ Soren Holm, “Goodbye to the Simple Solutions: The Second Phase of Priority Setting in Health Care,” 317, no. 7164 (1998).

²⁸⁰ Schwartzman, “The Completeness of Public Reason.”

that some will be stuck with policies that they can agree are reasonable – based on public, mutually acceptable, or relevant reasons, as they have variously been called – but think are less so than the ones that they support. Let us address the case of inconclusiveness first.

As captured succinctly in an article titled, “Establishing a fair process for priority setting is easier than agreeing on principles,”²⁸¹ Daniels and Sabin argue that when we cannot agree on substantive principles, we may nevertheless be able to agree on a fair procedure:

In the absence of a broadly accepted consensus on principles for fair distribution, the problem of fair allocation becomes one of procedural justice. The basic idea behind this appeal to procedural justice is quite familiar. When we lack consensus on principles that tell us what is fair, or even when we have general principles but are burdened by reasonable disagreements about how they apply, we may nevertheless find a process or procedure that most can accept as fair to those who are affected by such decisions. That fair process then determines for us what counts as a fair outcome.²⁸²

The procedural approach that they advocate for is captured by the four conditions of accountability for reasonableness.

These four conditions are:

- 1) *publicity*: decisions and their rationales must be fully transparent and publicly accessible,
- 2) *relevance*: rationales should appeal to “evidence, reasons, and principles that are accepted as relevant by fair-minded people who are disposed to finding mutually justifiable terms of cooperation,”
- 3) *revision and appeals*: there must be mechanisms for the challenge and revision of decisions, especially in light of new evidence or arguments,
- 4) *enforcement*: there must be assurances that the preceding three conditions are met.²⁸³

²⁸¹ Norman Daniels, “Accountability for Reasonableness: Establishing a Fair Process for Priority Setting Is Easier than Agreeing on Principles,” *BMJ* 321, no. 7272 (2000).

²⁸² Daniels and Sabin, *Setting Limits Fairly: Can We Learn to Share Medical Resources?*, 4.

²⁸³ *Ibid.*, 12.

It is important to note the way in which this account of procedural justice is not, in fact, fully procedural. Although the conditions of publicity, revision and appeals, and enforcement are purely procedural in nature, the practical application of the relevancy requirement leads us back to the problems associated with the ‘reasonable’ – ‘relevancy’ clearly displays public reason roots. So although it is claimed as an appeal to procedural justice, the relevance condition suggests the same kind of substantive constraints on legitimate reasons. This is acceptable for Daniels and Sabin because they are optimistic that we will be able to find reasons that all can agree are relevant in the resource allocation context. I am less optimistic about this, as the discussion in the previous section suggests, particularly when we move to other contexts like reproductive technologies. In other words, in contrast to Daniels and Sabin, I think that decisions will often need to be made in which *not* all can agree on which evidence, reasons, and principles are relevant.

Although they do not seem to think that the application of the relevancy requirement will be indeterminate then, they do admit that it may be inconclusive. Specifically, they consider what to do when public reasoning runs out, in which “a deliberation appeals only to reasons that all can recognize as acceptable or relevant, but that consensus about an outcome is still not achieved.”²⁸⁴ In such cases, because deliberation must inevitably be cut short, Daniels and Sabin suggest that, “[t]o settle the practical matter we rely on a majority vote.”²⁸⁵ Again it is important to remember that they find this vote acceptable because the policy options, while multiple, all rest on public reasons. Indeed, they justify majority rule in these instances because “the minority is not being compelled to do something for reasons it thinks irrelevant or inappropriate –

²⁸⁴ Ibid., 36.

²⁸⁵ Ibid.

even if it does not accept the weight or balance given to various considerations by the majority.”²⁸⁶

In the face of disagreement about the most ‘reasonable’ of a variety of options – in other words, when the relevancy requirement is inconclusive – the three other truly procedural conditions set up a framework or ‘rules of the game’ that make the majority vote and the ultimate reliance on a disputed balance of values more justifiable than it might otherwise be. How so? On the one hand, the conditions of publicity, revisions/appeal, and enforcement do indeed seem to construct a process that we can all agree is fair. This ability to reach consensus on procedural requirements is normatively important, as a way of respecting political autonomy in the tradition of social contract theories, as well as practically important, as a way of sustaining political support for the process. On the other hand, the conditions of publicity and revision/appeals crucially allow those who ‘lose out’ after the majority vote to see exactly what reasons and evidence were used to overrule their preferred policy option, and provides them with clearly accessible means of challenging that decision. Jane Mansbridge has compellingly argued that we must design our institutions to foster “informal deliberative enclaves of resistance,” on the understanding that “none of the criteria for legitimating coercion can produce in cases of ongoing disagreement an incontestably legitimate result.”²⁸⁷ I consider that true transparency about the basis of decisions, an avenue for appeals, and the prospect of future revision are necessary measures in the face of persistent disagreement to enable such deliberative resistance and thus treat those in the deliberative minority fairly.

²⁸⁶ Ibid.

²⁸⁷ Jane Mansbridge, “Using Power/Fighting Power: The Polity,” in *Democracy and Difference: Contesting the Boundaries of the Political*, ed. Seyla Benhabib (Princeton, N.J: Princeton University Press, 1996): 46, 54.

Importantly, based on these two defences of resorting to a fair process, the procedural requirements outlined by Daniels and Sabin will be useful in cases of indeterminacy just as much as in cases of inconclusiveness. Let us now, then, include these ‘harder’ cases of deliberations in which public reasoning is indeterminate. For instance, although clearly more sceptical about the ability to reach agreement on relevant reasons in the context of emerging biotechnologies, the Nuffield Council on Bioethics’ public ethics approach follows a very similar logic about the retreat to fair procedural conditions. The report recommends a set of ‘procedural virtues’ noting, “Rather than concrete prescriptions, for emerging biotechnologies we must...look to *how the business of policymaking and governance is carried out*, rather than its substantive content.”²⁸⁸

These procedural virtues are:

1. *Openness and inclusion*: Availability of information and access, where appropriate, for citizens to participate in biotechnology governance.
2. *Accountability*: Explicit acknowledgement of where responsibility for decisions lies and how actors may thus seek to influence or revise them.
3. *Public reasoning*: Cultivation of reasoning orientated to discovery of common ground and impartial interpretation of evidence.
4. *Candour*: Truthful representation of any uncertainties.
5. *Enablement*: Supporting wider political debate and proposals of alternative positions in public discourse.
6. *Caution*: Gathering deeper and more extensive knowledge before making policy decisions when there is a greater degree of uncertainty and ambiguity.²⁸⁹

Like the three procedural conditions of accountability for reasonableness, these ‘virtues’ – particularly openness and inclusion, accountability (note how it supports revision and appeals), candour, and enablement – set up fair rules of the game and provide citizens with the resources to challenge contested decisions. In particular, the virtue of enablement is closely aligned with the goal of fostering deliberative resistance, and

²⁸⁸ Nuffield Council on Bioethics, Chapter 4: Public Ethics and the Governance of Emerging Biotechnologies, “Emerging Biotechnologies: Technology, Choice and the Public Good.”

²⁸⁹ Ibid.

suggests how an arm's length body governing emerging biotechnologies may have a duty – as a matter of fair process – to actively educate the public about the state of the art, beyond any narrow consultation for the purpose of making specific decisions. Note also how the virtue of 'public reasoning' is much more relaxed than the relevancy requirement of accountability for reasonableness; it suggests that we should try to meet the demands of reciprocity but that mutually acceptable reasons may not always be forthcoming.

As a step before voting in cases where public reason does not yield a single outcome, Gutmann and Thompson²⁹⁰ and others²⁹¹ have suggested the necessity of striving for 'moral accommodation' in the face of deliberative disagreement. Due to the demands of reciprocity, deliberators should recognize the significance of having to base a political decision on one ordering of a set of mutually acceptable reasons, even when another reasonable ordering admittedly exists (inconclusiveness), or on the basis of reasons that others can reasonably reject (indeterminacy). They should recognize that such a decision, even if necessary, is non-ideal, and that the duty of reciprocity does not end simply because we reach the realm of deliberative disagreement.

As such, rather than immediately retreating to a majority rule solution, in which individuals base their voting on nonpublic considerations, deliberators should seek an 'economy of moral disagreement.'²⁹² This economizing requires that citizens both recognize the moral character of opponents' deeply-held beliefs and seek policies that as best as possible respond to and accommodate those beliefs. They should aim to negotiate a moral compromise, in which groups are willing to agree on a policy outcome even if it

²⁹⁰ Gutmann and Thompson, *Democracy and Disagreement*, 79-94.

²⁹¹ See Schwartzman, "The Completeness of Public Reason"; David B. Wong, "Coping with Moral Conflict and Ambiguity," *Ethics* 102, no. 4 (1992); George Sher, "Subsidized Abortion: Moral Rights and Moral Compromise," *Philosophy & Public Affairs* 10, no. 4 (1981); Frances M. Kamm, "The Philosopher as Insider and Outsider," *Journal of Medicine and Philosophy* 15, no. 4 (1990).

²⁹² Gutmann and Thompson, *Democracy and Disagreement*, 85.

is not either side's first preference. For an example of such an economy of moral disagreement, Gutmann and Thompson cite the Supreme Court's decision in *Planned Parenthood v. Casey*, which upheld a woman's general right to obtain an abortion while simultaneously recognizing the states' compelling interest in the potential life of the foetus that may justify certain restrictions, such as 24-hour waiting periods, provided they do not constitute an 'undue burden.'²⁹³

However, if a moral compromise cannot be reached, deliberators must resort to some other decision procedure. In the case of indeterminacy, Schwartzman – as a staunch public reason liberal – suggests that to avoid relying on nonpublic reasons we must resort to a lottery, since there are no public reasons on which deliberators can base their vote: “[s]ome additional consideration would be needed to tip the scales, which is why it is tempting to appeal to nonpublic reasons.”²⁹⁴ Instituting a lottery or a coin flip for arm's length decision-making seems implausible and practically unsustainable as a matter of public policy. A vote in which every deliberator counts equally seems just as procedurally fair at this point as a random procedure. Indeed, although voters would have to rely on nonpublic reasons in the end, we may also expect it to lead to epistemically better decisions on the whole than a coin flip or lottery. As Seyla Benhabib puts this appeal to the rationality drawn from Habermasian discourse ethics:

[I]n many instances the majority rule is a fair and rational decision procedure, not because legitimacy resides in numbers but because if a majority of people are convinced at one point on the basis of reasons formulated as closely as possible as a result of a process of discursive deliberation that conclusion A is the right thing to do, then this conclusion can remain valid until challenged by good reasons by some other group.²⁹⁵

²⁹³ Ibid., 87.

²⁹⁴ Schwartzman, “The Completeness of Public Reason,” 211.

²⁹⁵ Benhabib, “Toward a Deliberative Model of Democratic Legitimacy,” 72.

In sum, when arm's length bodies face the persistence of deliberative disagreement, when, in other words, appeal to public, mutually acceptable, or relevant reasons is inconclusive or indeterminate, they may attempt a moral compromise. Barring that, they may legitimately resort to a majority vote in which members rely on the reasons they find most persuasive after this process of reciprocal deliberation, so long as strong procedural conditions are in place. These conditions should be geared towards fostering the means of deliberative resistance for the 'losers' who must abide by decisions that they view as either less reasonable than an alternative or unreasonable. This means that they must have 1) the avenues to challenge these decisions, through an appeals process and/or regular revision of rules, and 2) the information to challenge these decisions, which requires a clear statement of the substantive basis of contested decisions and candid and balanced public information and educative outreach on the relevant issues. I turn now to incorporating these conclusions and others from the chapter into a concise framework for public accountability in the arm's length regulation of emerging biotechnologies.

7. A framework of public accountability as reciprocity for the arm's length regulation of emerging biotechnologies

This chapter set out to address the legitimacy problem that arm's length bodies face in light of the tension, illustrated in Chapter I, between the need to be politically independent and the need to be publicly accountable. A lack of guidance about how to navigate this tension has led to problems and confusion in the practical decision-making of arm's length bodies governing emerging biotechnologies, specifically the HFEA, NICE, and MHRA. In this chapter, I began by suggesting that although arm's length bodies by definition may not be held accountable through traditional lines of democratic

control, public accountability within the arm's length context may be conceptualized as careful consideration of the 'public interest' or 'common good.' It is to the end of understanding what precisely we mean by these terms that I have drawn herein on theories of public reason and deliberative democracy. I attempt to draw together my conclusions, about the important role of the criterion of reciprocity, the limits of public reason, and the necessity of procedural requirements, in a framework meant to answer the initial question of what it means for an arm's length body to be publicly accountable.

The four conditions of my framework, that of 1) reasoning reciprocally, 2) accepting deliberative disagreement, 3) voting, and 4) fostering the means of resistance, are all based in some sense on a commitment to the criterion of reciprocity. In this sense, my take on public accountability does require endorsement of a specific *substantive* goal or motivation of arm's length bodies' reasoning. It is not procedural "all the way down," but I take this to be neither possible nor desirable in an account of legitimacy.²⁹⁶ Rather, accepting the motivational force of reciprocity, I attempt to explain the kinds of *procedural* requirements that fall out of it. This allows critique and improvement of the way in which arm's length bodies operate, rather than the substantive content of their decisions.

(i) *Reasoning reciprocally*

First, and centrally, arm's length bodies regulating emerging biotechnologies in the UK ought to begin their deliberations with the goal of reasoning on the basis of reciprocity, rather than seeking to identify and implement an account of the 'whole truth,' even if they believe this to be an impartial or objective account. This condition has two key impacts on the processes of deliberation within arm's length bodies. First,

²⁹⁶ Frank I. Michelman, "How Can the People Ever Make the Law? A Critique of Deliberative Democracy," in *Deliberative Democracy: Essays on Reason and Politics*, ed. James Bohman and William Rehg (London: The MIT Press, 1997).

the demands of reciprocity require certain characteristics of deliberators, or in our terms, members appointed to arm's length bodies. They must be sincerely interested in finding reasons and other considerations (standards for evidence, for instance) that are mutually justifiable and must be willing to have their minds changed through deliberation, not necessarily on their first-order beliefs (those that make up their comprehensive doctrines), but on the set of public reasons and weight of public values on the basis of argument and empirical evidence. It is also desirable that they be adept at identifying, interrogating, and balancing principles or considerations, indicating that the inclusion of a practical philosopher with experience in moral and political theory would be valuable as a person well placed to guide the quite demanding requirements of reciprocity. Of course, there are certainly non-philosophers with these skills and experiences, too. And this should not be to the exclusion of others whose knowledge or skills will improve deliberation by, for instance, shedding light on types and standards of evidence or clinical realities, such as scientists, doctors, lawyers, or social scientists.

A second procedural requirement that follows from the commitment to reciprocity is the proper involvement of the public so as to elicit the range of public values on questions before the body. Public involvement allows the body to get an accurate sense of the landscape of public values: what, in other words, individuals with divergent comprehensive doctrines have to say on a given issue, like the permissibility of mitochondrial donation or whether medical conditions with no current treatment should receive some type of priority in technology appraisals. This helps address the concern that the values driving policy will become those of the 'elite' deliberators and detached from the broader public; as Alex Friedman worries in similarly calling for more public participation than exists in Daniels' and Sabin's accountability for reasonableness, "[H]ow can a process that by virtue of its very nature excludes almost everyone from

meaningful participation (from a very cynical point of view – everyone except the political philosophers who are proponents of such processes) confer democratic legitimacy on its outcomes?”²⁹⁷ Importantly, this involvement must tap into the values underlying public preferences: it is not enough to simply ask whether members of the public approve or disapprove of a novel biotechnology. Nor is it enough to know the *primary* normative commitment underlying members’ of the public preferences: it is important to realize, for instance, that those who think sex selection is wrong because it treats children inappropriately as commodities also recognize – as *relevant* – claims about reproductive autonomy and merely find that, in their mind, they are outweighed in this case.

This level of sensitivity is important in public involvement because deliberating members of an arm’s length body will thereafter have to subject these preferences to appropriate scrutiny to be sure they are 1) informed and 2) meet the requirement of treating others as free and equal, and then search for common, public values upon which to base policy. So for instance, drawing on an example expanded upon in the case study of the HFEA, members of the public might indeed oppose sex selection from the perspective of ‘what I think is morally right’ – they may think it the wrong way of building a family – but the arm’s length body must be responsible for ensuring that policy is based on ‘what we think can be justified on the basis of common values’ – and thus consider broader and overlapping considerations. The latter will depend upon the collection and proper interrogation of the former, but will not be dictated by it.

(ii) Accepting deliberative disagreement

Second, deliberators must be willing to recognize and accept the limits of public reason: that even when all reason reciprocally, there may nevertheless be deliberations

²⁹⁷ Alex Friedman, “Beyond Accountability for Reasonableness,” *Bioethics* 22, no. 2 (2008): 108.

that end in disagreement. Emerging biotechnologies, based on their common characteristics, may be expected to prove more impervious to public reasoning than the average political decision, whether by inconclusiveness or indeterminacy. In such situations, it may be tempting to seek a false consensus by over-circumscribing the realm of the ‘reasonable’ or ‘relevant.’ Deliberators must conscientiously avoid this temptation, and rather seek to candidly *flesh out* the deliberative disagreement that remains at the close of deliberation. This implicates the procedural requirement of transparency in a particularly strong form: the process of reasoning that an arm’s length body has gone through and the nature of any disagreement that reasonably persists must be documented fully in a publicly available explanation that accompanies a decision. This sharply contrasts with many policy documents that attempt to present a decision as a fully ordered and rationale consensus, in the process overlooking areas of reasonable disagreement.

(iii) *Voting*

Third, in such cases of deliberative disagreement, deliberators should first attempt to locate a moral compromise to which they can all agree. If this is not possible, they may legitimately turn to a majority vote to reach a decision. The basis of deliberators’ votes in this case will necessarily depend upon considerations that are not fully public in nature, but this is justifiable because decisions must be made even after an attempt to reason reciprocally has produced no mutually justifiable positions. Deliberators should at this point rely on the consideration they find to be the strongest. Procedurally, this means that a final consensus is not required of deliberators and that the public explanation document should include an account of the arguments that different deliberators found to be most convincing.

(iv) *Fostering the means of deliberative resistance*

Fourth, because decisions may be based on reasons that a minority may reasonably reject, arm's length bodies have a duty to both 1) create avenues for challenging these decisions and 2) institutionally support the deliberative resistance of those who lose out. Procedurally, this means there must be institutional pathways in place to allow for the challenge and revision of decisions. It is particularly desirable to set a period after which the decision will be revisited or at least considered for revisitation in light of new arguments or evidence. The chance to revise decisions, whether through institutionally-set review periods or externally-brought appeals, ensures the public both that contested decisions are provisional and that there is timely reconsideration of issues based on the accumulation of more and better evidence. However, the valid grounds of appeal must be deliberative in nature: they must engage with the reasoning of the arm's length body on the understanding that reciprocity must guide any reconsideration of the decision.

To support such future deliberative resistance, a procedural condition of publicity is necessary and must mean significantly more than making the process of decision-making transparent, as has seemed to be the case in many common reforms of arm's length bodies that make meetings open to the public. The requirement that decision-making be public must not be understood solely as making the *process* of decision-making clear. It must also be understood as making the substantive basis of decisions clear, particularly when decisions must be made in the face of deliberative disagreement. This will require a clear statement of 1) how the body involved the public so as to elicit the values underlying a range of perspectives, 2) the considerations or reasons that the group decided were mutually acceptable or that all could agree were relevant, 3) how the

group deliberated about ordering and specifying those considerations, and 4) how the final decision was reached (consensus/vote).

Where other policy positions were considered justifiable on mutually acceptable reasons – perhaps most likely in cases of allocation – deliberators must attempt to articulate how they weighed or specified values in settling on the ultimate position. Where all policy options required appeal to nonpublic considerations – as may be expected in many decisions on reproductive technologies – deliberators must clearly state the contested moral principles or claims that ground the position. This strong interpretation of publicity is necessary to support the deliberative resistance of those who have lost out: it provides the process and grounds of a decision that those who disagree with it must be able to engage with and respond to in their challenges.

Finally, and returning the relationship between the arm's length body and the public, there must be a general programme of public outreach and education aimed at disseminating the idea of reciprocity as the core of public accountability, the body's processes and decisions, and the state of both scientific and ethical debates around the emerging biotechnologies it regulates. This will require being publicly recognizable and having a staff large enough to help carry out such a public outreach programme, particularly by working with the media. This public outreach, as opposed to public consultation, furthers the goal of informing the public of the basis of the rules that they must live by and thus contributes to fostering the means by which they may, if they find such rules out of step with their values, push back against them in a deliberative manner.

To the extent that arm's length bodies meet these four conditions, they operate legitimately and are properly accountable to the public. However, the way in which these conditions and their procedural requirements are specified in practice will depend upon

the institutional context in which they function; while it is a general framework, it does not result in a one-size-fits-all model. Rather, it must be carefully substantiated in each of the three arm's length bodies I consider if it is to help resolve the practical manifestation of the legitimacy problem that each faces.

Chapter IV. Reciprocal Reasoning in the HFEA: Implications for Public Consultation, Lay Representation, and Ethics Expertise

In Chapter II, I suggested that the practical manifestation of the legitimacy problem in the HFEA can be seen in accusations that the body pays either too much or too little attention to the views of the public in making controversial value judgements. The procedural method of ‘balancing’ reasons and interests from the general public and various stakeholders cannot explain, I noted, how members reach conclusions on these contested values through deliberation. Some substantive principle must be at work determining the ‘weight’ afforded to different reasons in this method of balancing. And, indeed, my previous empirical research found that this metaphor of balancing in fact masks a particular commitment to prioritising considerations of health and welfare over alternative values, like respect for the special status of the embryo or the proper way of valuing children as a parent.²⁹⁸ Regardless of the substantive principle(s) determining the weight of various considerations in a method of balancing, a key point to recognize is that the existence and content of such principles is obscured by the apparently neutral and purely procedural concept of balancing. So long as these principles, which guide the *actual* method of reasoning, are hidden, the HFEA may reasonably be critiqued for failing to make transparent – and to justify – its true approach to the process of decision-making.

²⁹⁸ Edwards, “Methods of Legitimation.”

Rectifying this problem requires thinking critically about what the basic goal or purpose of the HFEA's deliberations are, which see scientists, doctors, lawyers, philosophers and lay members reasoning together while drawing on the results of public consultations, and making this basic purpose clear. In this sense, the HFEA, relative to the other two case studies considered, presents the purest form of the legitimacy problem, understood as the tension between accountable and independent decision-making in arm's length bodies. The challenge here is to determine what a legitimate method of deliberative reasoning would look like in a body that is deeply committed both to public engagement and consultation, on the one hand, and informed, rigorous consideration of the ethics of reproductive technologies on the other. How, especially when the two sources seem to suggest conflicting policy decisions, can the HFEA attend to both the results of public consultation, as the attempt to be accountable, and the results of members' ethical reasoning, as the attempt to make decisions independently and rationally?

In this chapter, I apply the conditions of my public accountability framework to work towards the resolution of this challenge, focusing specifically on fleshing out the practical implications of the first condition of the framework, that of reasoning reciprocally. I begin by defending this goal of reciprocity in the context of an arm's length body making public decisions on ethics, as the HFEA does, against both 1) the goal of representing public views or values, and 2) the goal of identifying the truth of the matter, or the position most capable of withstanding philosophical challenge. If the former may be understood as the 'mirror model' of public ethics committees and the latter is understood as the 'critical model,' I propose a reciprocal model that overcomes a number of problems faced by the traditional pair. Next, I illustrate how this reciprocal model can provide a justified role for the frequently confused practices of public

consultation and the appointment of lay members and philosophers to arm's length bodies. Thus, in addition to providing a central method of reasoning, the reciprocal model can guide these processes common within arm's length bodies like the HFEA. Finally, I look at a 'hard' case for the HFEA, in which the results of public consultation conflicted with members' apparent deliberative conclusions, the case of non-medical sex selection. With reference to my accountability framework and the reciprocal model of ethics committees I have developed in this context, I suggest what went wrong in the process of decision-making about non-medical sex selection and how it might be improved.

A more general explanation of the following three chapters and their relationship to the framework set out in the previous chapter is warranted. These 'case study' chapters on the HFEA, NICE, and MHRA are primarily directed towards using the framework described in Chapter III to improve the legitimacy of decision-making by these three arm's length bodies. In this sense, they are concerned with the translation of theoretical claims into practice. Although I have attempted to outline the general requirements of public accountability in the arm's length context with the four conditions of my framework, it remains to be seen how these general requirements can and should be specified in addressing the real-world challenges facing each body. Put otherwise, how can these largely abstract conditions usefully impact practice? In the following three chapters, I articulate what follows from my general claims about public accountability for the resolution of the practical challenges set out in Chapter II in the three diverse policy making settings considered.

1. A third way: defending reciprocity against ‘representation’ and ‘reason’

The central role of public consultations in the HFEA starkly highlights the tension that it faces between ‘reason’ and ‘representation,’ which often (although not always) pull in opposite directions and suggest two very different methods of reasoning for an arm’s length body like the HFEA. As Caroline Mullen usefully frames this tension, a public decision on ethics may be understood as legitimate “either in the sense of legitimacy in virtue of reflecting the opinions of the public whose interests are affected, or in the sense of being able to withstand and respond to challenges presented in ethical debate.”²⁹⁹ These two understandings of legitimacy in public decision-making – based on 1) the accurate representation of citizens’ preferences, or 2) the strength of reasons and arguments – indicate two radically different potential goals of an ethics commission or licensing authority like the HFEA.

On the first understanding, the body could simply attempt to collect an accurate representation of public opinion on a given issue, through a representative poll or survey, and base decisions on what it found the majority of the public to prefer. Educational materials could precede such polling or surveying so that the expressed preferences might at least be more informed than those of the average member of the public, who may not have been exposed to or given much thought to the implications of new reproductive technologies. Alternatively, although still based on the idea of *reflecting* public perspectives, a body like the HFEA might focus more on identifying and basing decisions on existing public values rather than existing public opinion on specific policy positions. In this way, it might operate on what Simon Clarke has called the ‘mirror model’ of ethics committees:

²⁹⁹ Caroline Mullen, “Representation or Reason: Consulting the Public on the Ethics of Health Policy,” *Health Care Analysis* 16, no. 4 (2008): 397.

The committee is to reflect the values of the society of which it is a part. It makes decisions by applying ethical standards which it takes to be inherent in the culture of the society. On this view, ethical standards are a set of ‘common sense’ principles and ideals that committees should be guided by.³⁰⁰

On this approach, a national committee might similarly use polls and surveys, although representative focus groups might be the most useful method, to elicit the values underlying public preferences and use these values, and the magnitude with which they are represented in the general public, uncritically in making a decision.

These models of merely reflecting public opinion or even underlying public values in an *uncritical* manner appeal to legitimacy through accurate representation of the public’s beliefs and values, and the reflection of these prevailing opinions or morals in policy. They ask, in a sense, for an arm’s length body like the HFEA to behave institutionally more like a direct referendum or Parliament, to behave in a more *democratically* accountable manner. But, as explained in the previous chapter, a large part of the reason for setting decisions about the permissibility of reproductive technologies at arm’s length is to circumvent such aggregative democratic institutions by allowing thoughtful deliberation about and scrutiny of the *reasons* for different policy options: to avoid what Eve Garrard and Stephen Wilkinson have called “the uncritical identification of ‘might’ (majority or social convention) with ‘right’ (what people really ought to do).”³⁰¹ This deliberative alternative to the aggregative model of democracy allows us to protect basic rights and liberties and ensure more generally that decision-making considers the ‘*public* interest,’ not merely the current majority’s interest. Indeed, the impulse even within the mirror model to ensure that expressed preferences in polls or

³⁰⁰ Simon Clarke, “Two Models of Ethics Committees,” *Journal of Bioethical Inquiry* 2, no. 1 (2005): 41.

³⁰¹ Stephen Wilkinson and Eve Garrard, “Mind the Gap: The Use of Empirical Evidence in Bioethics,” in *Bioethics and Social Reality*, ed. Matti Häyry, Tuija Takala, and Peter Herissone-Kelly, Value Inquiry Book Series 165 (New York: Rodopi bv, 2005), 83.

surveys are *informed* reveals the intuition that decision-making should meet at least some standards of rationality rather than mere aggregation. On these bases, we can fairly easily dismiss the notion that the HFEA's purpose should be to simply mirror prevailing public opinion or the most widely-held values in policy.

The second account of legitimacy in public decisions on ethics – based on the degree to which reasons and arguments underlying decisions may withstand philosophical challenge – implicates a contrasting model of ethics committees, which Clarke terms the 'critical model.' In this model, "[A]n ethics committee is to take ethical standards held in society but subject them to critical evaluation...Beliefs may be founded on unjustified assumptions, and/or may be inconsistent with other beliefs that have rational foundation."³⁰² This, indeed, seems to be the general model to which the HFEA adheres. In addition to my own study of the HFEA's method of balancing, which illustrates the evaluative weighing of considerations drawn from public consultation, a set of interviews with the HFEA carried out by political scientist Alfred Moore led him, too, to conclude that the HFEA sees its role as scrutinizing the "views and positions gathered in the consultations [that] constitute raw data" in order to "distinguish the 'real' ethical considerations from simple prejudice."³⁰³ Unlike the mirror model, then, the critical model allows for important reflection on the *strength* of the arguments underlying policy positions; it refuses to accept the opinions or values of the public at face value, instead testing these 'raw' considerations in deliberation.

But to what test, exactly, are these views or values put? Asked otherwise, on what basis or metric does a 'critical' ethics committee identify the "real" ethical considerations and assign more or less weight to them? This model seeks as its

³⁰² Clarke, "Two Models of Ethics Committees," 41.

³⁰³ Moore, "Public Bioethics and Public Engagement," 207.

overarching goal the ‘right’ or ‘best’ answer to a policy dilemma, where this answer is understood as the one most capable of withstanding philosophical challenge or ethical debate. With such a goal, deliberators will be directed to appeal to what they see as the best reasons or strongest arguments, *all things considered*. And the ethical debate to which such arguments must respond will be that set by the other deliberators and *their* particular conceptions of the best arguments, all things considered. To the end of identifying an objective truth of the matter, and with a lack of constraints on their method of reasoning, deliberators may appeal to considerations drawn from disputed religious, moral or philosophical comprehensive doctrines, which will stand or fall in deliberations based on the extent to which other participants are like-minded.

To be sure, attempting to create a diverse committee whose members hold various moral perspectives and are willing to have their minds changed will go some way to preventing the outcomes of deliberations from being ‘fixed’ by who gets a seat at the table. If the committee’s composition is ‘engineered’ or ‘gerrymandered,’ as some have worried the presidential bioethics commissions in the United States tend to be,³⁰⁴ the concern is that the ‘strongest’ arguments will be predetermined by the choice of members and the contested values or philosophical approaches that they hold. The idea is that if a conservative politician appoints a bioethics commission packed with conservative bioethicists, we should be unsurprised if the outcome of so-called ‘critical’ deliberation determines conservative values or arguments to be the ‘strongest.’ Equally, those who do not share these values will have little reason to recognize the committee’s recommendations as being in fact based on the strongest arguments, because the structure of the deliberative space has already ruled out introduction of a range of

³⁰⁴ Lisa M. Rasmussen, “Engineering, Gerrymandering and Expertise in Public Bioethics,” *HEC Forum* 18, no. 2 (2006); Ronald M. Green, “For Richer or Poorer? Evaluating the President’s Council on Bioethics,” *HEC Forum* 18, no. 2 (2006).

arguments and values, specifically those that they find most compelling. It is, in Habermasian terms, a very non-ideal speech situation, one in which we have little reason to believe that the outcomes of the process of deliberation may claim to be epistemically better or more rational by virtue of being subjected to deliberative scrutiny.

Suppose, then, that we do not gerrymander the composition of the committee. How does the critical model recommend that we select members, beyond seeking to have a range of ethical outlooks and people who are willing to follow the ‘force of the better argument’? The ultimate goal, recall, of the critical model is to produce answers most capable of withstanding ethical debate. To be clear, the underlying assumptions here are morally realist and objective rather than irrealist, subjectivist, or relativist (in that there are objectively better or right answers to ethical questions), and not completely epistemically sceptical (in that we can come to know these better answers through practical reasoning and so make moral progress). Some will reject these two assumptions and may therefore dismiss the critical model out of hand. I think both assumptions are true, but will not attempt an argument for them here. Rather, sidelining these longstanding metaethical and epistemic debates, let us accept for the sake of argument both assumptions in order to consider what would follow.

If reaching the moral truth of the matter is the goal, we ought it seems to empower those who may be relatively better at reasoning practically about what we ought to do: those who know the facts or evidence on the issue at hand, who are familiar with moral concepts, who may bring together evidence with a moral position or method of moral reasoning, who can clearly construct sound arguments with attention to consistency and the potential for bias, and who have the time to consider ethical questions. These criteria are drawn from Peter Singer’s account of what constitutes

‘moral expertise,’³⁰⁵ a much-disputed concept.³⁰⁶ Taking this argument for the critical model to its extreme, it is unclear on what grounds such a committee – with members skilled in ethical reasoning – would need to consult or engage with the public at all, other than perhaps to gain more or better information about the facts of an issue or the expected consequences of certain courses of action.

Related to the metaethical and epistemic debates, I do not wish to dive too deep into the questions of whether moral expertise exists and what it consists of. However, it is worth saying two things about what the exercise of this kind of moral expertise would mean in the context of policy making. First, the conviction about what ‘moral position or method of moral reasoning’ – from Singer’s criteria – to use is precisely the question at issue; we can easily agree, I think, that individuals with more information and time to consider an issue, who are conscious of bias and consistency, and who can construct valid arguments whose conclusions follow from the premises will, on the whole, reason better about an issue at hand than those who do not. However, so long as the values or moral framework upon which the conclusion is premised are deeply disputed, we run into a problem similar to that in the gerrymandering case: that those who disagree with the moral premise have little or no reason to think the decision is a good one, even if they grant that these other virtues of the ‘moral expert’ will tend to lead to better outcomes in a deliberative process. But, again for the sake of argument, assume that such a committee *is* able to identify and empower moral experts, who are better at judging what we ought to do, and by virtue of this *is* more likely to ‘get it right,’ regardless of

³⁰⁵ Peter Singer, “Moral Experts,” *Analysis* 32, no. 4 (1972).

³⁰⁶ See Jan Crosthwaite, “Moral Expertise: A Problem in the Professional Ethics of Professional Ethicists,” *Bioethics* 9, no. 5 (1995); Christopher Cowley, “A New Rejection of Moral Expertise,” *Medicine, Health Care and Philosophy* 8, no. 3 (2005); Bernward Gesang, “Are Moral Philosophers Moral Experts?,” *Bioethics* 24, no. 4 (2010); David Archard, “Why Moral Philosophers Are Not and Should Not Be Moral Experts,” *Bioethics* 25, no. 3 (2011).

what anyone who disagrees with their decision thinks. Do we then have reason to accept the critical model of ethics committees and the legitimacy of their decisions?

This brings to the fore arguments for epistocracy: the idea that those who know best ought to rule, with the most famous account being that of Plato's philosopher-kings. In rejecting epistocracy, David Estlund usefully refers to what he calls the expert/boss fallacy, the idea that authority does not flow *simply* from expertise:

Even if we grant that there are better and worse political decisions (which I think we must), and that some people know better what should be done than others (we all think some are much worse than others), it simply does not follow from their expertise that they have authority over us, or that they ought to.³⁰⁷

The type of principle that “blocks,” in Estlund's terms,³⁰⁸ the inference from expertise to authority is precisely that captured by the criterion of reciprocity, what I have set out as the central condition of public accountability in the arm's length context. As argued in Chapter III, recognizing the burdens of judgement and the resulting reasonable pluralism of moral comprehensive doctrines, we must try to give others reasons for coercive political decisions that they, holding divergent but reasonable doctrines, can accept, regardless of whether we find their views true or right. The Warnock Committee, which proposed the creation of the HFEA, seemed to recognize these limits on the exercise of moral authority:

Members of the Inquiry were reluctant to appear to dictate on matters of morals to the public at large. They were also keenly aware that *no expression of their own feelings would be a credible basis for recommendations, even if they all felt exactly alike...* If, as we believe, it was our task to attempt to discover the public good, in the widest sense, and to make recommendations in light of that, then we had, in the words

³⁰⁷ David M. Estlund, *Democratic Authority: A Philosophical Framework* (Princeton: Princeton University Press, 2009): 3.

³⁰⁸ *Ibid.*, 40.

of one philosopher, to adopt ‘a steady and general point of view’
(emphasis added).³⁰⁹

In other words, the Committee of Inquiry recognized that it could not justify a public position on the permissibility of IVF and its associated practices on the basis that *they* found it to be based on the most convincing arguments, even if they could come to unanimous agreement on this ‘right’ answer. Their concept of the adoption of a ‘steady and general point of view’ is not well elaborated in the report. However, on my account, the criterion of reciprocity is meant to provide and specify this perspective.

Now we can see why the critical model of ethics committees is misguided. In the pursuit of the answer most capable of withstanding philosophical challenge, it asks a group of elite deliberators to base public policy on the truth as they see it, all things considered. Even if we are convinced, based on the expertise of the deliberators, that they will in fact produce the right answer, or at least better ones, to the practical ethical problems before them (an appeal to the epistemic virtue of the model), this way of reasoning and justifying collective decisions does not meet the demands of what we owe other citizens when we recognize them as free and equal. Rather than values or principles drawn from comprehensive doctrines, we must make every effort to limit the reasons we give for coercive political decisions to those that are public or mutually acceptable. The critical model thus makes the mistake of prioritizing the identification of an objectively true answer above the identification of a *mutually justifiable* one.

Remember, of course, that such reasons need not actually be accepted by everyone; they must rather be ones that others who seek fair terms of cooperation and recognize people as free and equal could accept. This does not, then, result in a

³⁰⁹ “Report of the Committee of Inquiry into Human Fertilisation and Embryology,” 1.

relativistic ‘anything goes’ account of mutual acceptability, for, as Joshua Cohen explains, if one accepts at the beginning of deliberation that others are free and equal,

...then one cannot accept as a reason within that process that some are worth less than others or that the interests of one group are to count for less than those of others. And these constraints on reasons will limit the substantive outcomes of the process; they supplement the limits set by the generic idea of a fair procedure of reason giving.³¹⁰

The criterion of reciprocity as the fundamental condition of a deliberative *procedure* is thus capable of ensuring the protection of certain *substantive* liberties; we need not seek reasons that would be acceptable to a citizen who holds, for example, racist or bigoted beliefs.

In arguing against the mere aggregation of preferences in policy making, Jon Elster famously contends that collective decisions require the behaviour of the forum rather than the market; but we may also usefully distinguish between the behaviour of the forum and the behaviour of the philosophy academy. Just as the goal of a body like the HFEA should not be the mere aggregation of preferences implied by the mirror model, it should not be the search for the comprehensive moral truth of the matter implied by the unconstrained critical model. My framework for public accountability in the arm’s length context suggests a third way based on the criterion of reciprocity, which I shall refer to as the ‘reciprocal model’ of ethics committees. On this account, the HFEA ought to look for set of values on which to base decisions that all citizens who recognize others as free and equal and understand the political domain as an opportunity to seek fair terms of cooperation could accept. In the event that such reasons are not forthcoming, because of fundamental disagreement on the status of the embryo or the proper value of the family, the committee ought to search for a moral compromise. Only

³¹⁰ Cohen, “Deliberation and Democratic Legitimacy,” 413.

if these two stages fail to yield a determinate outcome should HFEA members resort to a vote, in which deliberators' votes will be based – justifiably – on the arguments that members find to be most convincing.

2. Implications of the reciprocal model for procedural requirements: public consultation and appointment of deliberators

What does this reciprocal model mean for procedural requirements, like the composition of the body and the practice of public involvement, in relation to the mirror and critical model? On the one hand, it means that public involvement has a secure place in the operation of committees like the HFEA that make public decisions on ethics, albeit a carefully tailored one. This is in contrast to the critical model, where public involvement may be superfluous or merely an act of political choreography³¹¹ to avoid appearing disconnected from the public sphere. Equally, however, the reciprocal model maintains that while the mirror model is right to insist on the involvement of the public, it has misinterpreted the role of such involvement and the way in which the public should be *situated* in relation to the arm's length body. Public involvement must be more deliberately designed with its purpose in mind than it has tended to be in the Patient and Public Involvement movement thus far.³¹²

First, we must make the basic distinction between involvement – either of the general public or specific patient groups – for the purpose of standard *meeting* and for

³¹¹ The notion of public consultation as political choreography comes from Graeme Cheeseman and Hugh Smith, "Public Consultation or Political Choreography? The Howard Government's Quest for Community Views on Defence Policy," *Australian Journal of International Affairs* 55, no. 1 (2001).

³¹² For critiques of public consultation as part of the larger PPI movement, see especially Jonathan Ives, Sarah Damery, and Sabi Redwod, "PPI, Paradoxes and Plato: Who's Sailing the Ship?," *Journal of Medical Ethics* 39, no. 3 (2013); Dominique Florin and Jennifer Dixon, "Public Involvement in Health Care," *BMJ* 328, no. 7432 (2004).

the purpose of standard *setting*.³¹³ In discussing the method of reasoning and the values upon which policies should be based, I have not been concerned with the first purpose of consultation, that of standard meeting: I assume it to be fairly uncontroversial that if the HFEA has set as a goal, for instance, an increase in the number of egg and sperm donors in the UK to reduce waiting times for infertile couples, they may reasonably consult the public, as potential but largely untapped donors, in order to determine how that goal might be achieved. It is the second purpose of consultation for standard setting, for determining the *values* upon which policies are based, that the reciprocal model is concerned with tailoring. There will be some back and forth on this obviously: it may be that the most efficient way to increase donors is to allow large commercial payment in exchange for sperm and eggs, but whether these means are justifiable even for an already accepted end may still be questioned.

Public consultation for standard setting, I have suggested, must seek to elicit the range of reasons and arguments that fair-minded citizens recognize as relevant in answering the ethical question at stake. In particular, public consultation of this type does not serve the role of a representative public opinion poll, and so it should not be constructed or carried out as such. The fact that the majority of individuals consulted register a preference for limiting IVF to married couples, or suggest – to show that this holds as much for values as preferences – that the value of a stable two-parent family is the overriding one in a decision about who should have access to IVF, provides very little useful information for a reciprocal ethics committee. As Garrard and Wilkinson incisively note, “If we are doing ethics (as opposed to merely trying to arrive at a political compromise) then the arguments (or lack thereof) must surely prevail over the

³¹³ The distinction between standard setting and standard using (which predicts the concept of standard meeting) was discussed in Chapter III and comes from Majone, “Europe’s ‘Democratic Deficit,’” 5-7.

views of the 80%,” where the arguments for this majority view are unsound, “and *this makes one wonder why we bothered to spend money on the market research in the first place.*”³¹⁴ While they and others speak of unsound arguments where I would speak of arguments not based on mutually acceptable reasons, the futility of the mere tallying-up style of public opinion polling is the same.

For the product of public consultation to be fit for purpose, namely use by members of the HFEA in reasoning reciprocally to a justifiable decision, it needs to clearly outline the full range of reasons that diverse citizens recognize as relevant, not merely what individuals see as their personal values or moral commitments. This suggests the desirability of focus groups or deliberative polling methods that allow the probing of citizens’ views and values on both a specific question, like the permissibility of mitochondrial donation, and the broader context in which it fits, like when a technique is ‘safe enough’ for use or in which circumstances embryo selection or manipulation is permissible. I will return to this question of what good public consultation should look like in the final section, which takes as an example the HFEA’s policy reviews of sex selection.

Beyond the practice of public consultation, the reciprocal model also has important things to say about the *composition* of the HFEA, particularly the phenomena 1) of appointing philosophers to bodies like the HFEA, and 2) of appointing ‘lay’ members to such bodies. Philosophers are not compulsory, though they are common on such bodies, but there tends to be an explicit requirement or quota for lay members. Meanwhile, philosophers appointed to the HFEA are simultaneously labelled as *lay* members and as members with *expertise* in philosophy,³¹⁵ making the traditional reading

³¹⁴ Wilkinson and Garrard, “Mind the Gap: The Use of Empirical Evidence in Bioethics,” 87.

³¹⁵ See, for example, Human Fertilisation and Embryology Authority, “Annual Report and Accounts 2009/10,” July 14, 2010, http://www.hfea.gov.uk/docs/Annual_Report_2010_Web_reduced.pdf, 25.

of lay as “not having professional qualifications or expertise” confusing. This naturally leads us to ask what exactly is meant by *lay* in this context, and what we hope lay members – in virtue of their laity – will bring to deliberations. One justification for lay membership in health governance cites the special information or knowledge that lay members possess that traditional medical or scientific ‘experts’ in this area do not, the seemingly paradoxical idea of ‘lay expertise’³¹⁶; a second stresses the value of the ‘disinterested’ perspective of lay members relative to the experts³¹⁷ or stakeholders.³¹⁸ More prolific than such specific justifications are over-simplistic claims that lay members make deliberation or decisions more representative, accountable, and/or democratic. In fact, there seems to be a common recognition in the literature both that lay representation on health authorities is becoming the norm and that substantial confusion still persists about what role lay members play.³¹⁹ Despite the explosion of lay members on national and local, advisory and decision-making, research and policy committees alike, “There is little empirical research on their role and no thought-out strategy about how lay people will be recruited, what skills they need or how they should be held accountable.”³²⁰

Similarly, we might wonder what role philosophers are meant to play in this context. As with lay members, it is well established that philosophers and ethicists have

³¹⁶ Lindsay Prior, “Belief, Knowledge and Expertise: The Emergence of the Lay Expert in Medical Sociology,” *Sociology of Health & Illness* 25, no. 3 (2003); Steven Epstein, “The Construction of Lay Expertise: AIDS Activism and the Forging of Credibility in the Reform of Clinical Trials,” *Science, Technology & Human Values* 20, no. 4 (1995).

³¹⁷ Robert Evans and Alexandra Plows, “Listening Without Prejudice? Re-Discovering the Value of the Disinterested Citizen,” *Social Studies of Science* 37, no. 6 (2007).

³¹⁸ See the idea of the ‘pure public’ rather than the ‘affected public’ or ‘partisan public’ in Kathrin Braun and Susanne Schultz, “... a Certain Amount of Engineering Involved’: Constructing the Public in Participatory Governance Arrangements,” *Public Understanding of Science* 19, no. 4 (2010).

³¹⁹ See Cathy Charles and Suzanne DeMaio, “Lay Participation in Health Care Decision Making: A Conceptual Framework,” *Journal of Health Politics, Policy and Law* 18, no. 4 (1993) and, more recently, Mark B. Brown, “Survey Article: Citizen Panels and the Concept of Representation,” *Journal of Political Philosophy* 14, no. 2 (2006).

³²⁰ Christine Hogg and Charlotte Williamson, “Whose Interests Do Lay People Represent? Towards an Understanding of the Role of Lay People as Members of Committees,” *Health Expectations: An International Journal of Public Participation in Health Care and Health Policy* 4, no. 1 (2001): 3.

increased in prominence in governance arrangements even though their status and role is unclear: “Notwithstanding the significance of the role of ethicists in contemporary public life, the concept of ethical expertise has been, and continues to be controversial,” note three such ethicists in surveying the current landscape.³²¹ David Archard, himself a philosopher serving on the HFEA, has posed the question, “[W]hy would moral philosophers be asked to serve on these bodies if it was not thought that they possessed some form of special expertise?”³²² But do they actually possess a special kind of expertise and, if so, what form does it take?

Let us begin with a fuller account of the practice of appointments in the HFEA than was provided in Chapter II. Appointments are made by the Secretary of State for Health with significant input from the chair and chief executive of the HFEA, but they are subject to the Code of Practice for Ministerial Appointments to Public Bodies.³²³ This code is based on the recommendations of the first report of the Committee on Standards in Public Life, commonly referred to as the Nolan Report or Nolan principles for its chair, Lord Michael Nolan. The Committee, formed in 1994 to address widespread concerns about the loss of public trust in government, proposed that quango appointments be made on ‘merit’ alone and that a new, independent Commissioner for Public Appointments audit these choices to ensure that they are not made on the basis of personal or political patronage.³²⁴ This sought to allay public anxieties at the time about the lack of transparency in the appointments process and the belief that “appointments

³²¹ Norbert L. Steinkamp, Bert Gordijn, and Henk A. M. J. ten Have, “Debating Ethical Expertise,” *Kennedy Institute of Ethics Journal* 18, no. 2 (2008): 173.

³²² Archard, “Why Moral Philosophers Are Not and Should Not Be Moral Experts,” 120.

³²³ Commissioner for Public Appointments, “Code of Practice for Ministerial Appointments to Public Bodies” (London, April 1, 2012), <http://publicappointmentscommissioner.independent.gov.uk/wp-content/uploads/2012/02/Code-of-Practice-20121.pdf>.

³²⁴ Chairman Lord Nolan, “Standards in Public Life: First Report of the Committee on Standards in Public Life,” Cm 2850-I, (May 1995), https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/336919/1stInquiryReport.pdf.

[were] being unduly influenced by party political considerations.”³²⁵ The move to consolidate merit-based appointments was founded at least in part, then, on the desire to block the type of gerrymandering of a committee tasked with making independent public decisions, which I discussed above. On what grounds of merit, however, should HFEA members be appointed? What, in other words, does merit mean in this context?

In proposing basic requirements for the future HFEA, the Warnock Committee noted, “The new body will need access to expert medical and scientific advice. We would therefore envisage a significant representation of scientific and medical interests among the membership.”³²⁶ Indeed, it seems uncontroversial that a body tasked with regulating any kind of biotechnology must be able to comprehensively grasp the scientific concepts behind the technology, be fluent in the available evidence on it and the standards of evidence in the field, and retain a firm connection to the clinical realities of the technology’s practical application, i.e. how it may be used and what it means for patients, particularly relative to other, currently available options. The appointment of scientific and medical experts thus ensures that decisions are well informed by including those who are most knowledgeable about the *facts* of the technology and the *social context* in which it is meant to function.

The Report immediately goes on to say, however, that:

...this is not exclusively, or even primarily, a medical or scientific body. It is concerned essentially with broader matters and with the protection of the public interest. If the public is to have confidence that this is an independent body, which is not to be unduly influenced by sectional interests, its membership must be wide-ranging and in particular the lay interests should be well represented.³²⁷

³²⁵ Ibid., 66.

³²⁶ “Report of the Committee of Inquiry into Human Fertilisation and Embryology,” 75-76.

³²⁷ Ibid., 76.

This passage illustrates well the initial impulse behind requiring lay membership on bodies that license new and controversial biotechnologies, like embryo research and various clinical applications of IVF. At the moment in which the Warnock Report was conceived, which some have linked to the birth of ‘public bioethics’ in the UK,³²⁸ the main concern was with removing sole authority for scientific or medical decision-making from scientists and doctors. The Report speaks of the Inquiry’s creation as an effort to address “a sense that events were moving too fast for their implications to be assimilated...unease at the apparently uncontrolled advance of science,” highlighting the concern that scientific and technological developments were progressing rapidly and without public oversight or scrutiny.³²⁹

The reaction to this sentiment of the unbridled pursuit of technological advancement was to formally situate ‘society’ alongside the medical profession and scientific researchers in the making of decisions about the desirability and permissibility of certain avenues of research or clinical practice. In the words of the Warnock Committee, “[S]cientists must not be unduly restricted in pursuing their research interests especially when this may produce direct therapeutic benefits. *But doctors and scientists work within the moral and legal framework determined by society.*”³³⁰ The move to create a statutory licensing authority – the HFEA – was thus a move to exert public control over a previously self-regulated area, to assume for society the authority to set the moral limits of scientific endeavours in the area of reproduction. In this sense, it represented a backlash against a certain kind of expert/boss inference, specifically that of the technical expert/moral boss: simply because scientists *know* more about the scientific or technical questions involved in reproductive technologies does not mean

³²⁸ See Wilson, “Creating the ‘ethics industry’”; Duncan Wilson, *The Making of British Bioethics* (Manchester: Manchester University Press, 2014).

³²⁹ “Report of the Committee of Inquiry into Human Fertilisation and Embryology.”, 4.

³³⁰ *Ibid.*, 75.

that they have any special authority to determine what the limits on the use of those technologies *ought* to be.³³¹

Thus, although doctors and scientists appointed to the HFEA would provide much needed advice – would draw, in other words, on their technical expertise – lay representation was to prevent decisions from becoming “unduly influenced” by them – to prevent their expertise from illegitimately shoring up special authority in such decisions. The Warnock Committee’s recommendation that there be “substantial lay representation...and that the chairman must be a lay person” thus sought to check the domination of decision-making by medical and scientific interests. More recent statements have suggested the endurance of this defence of lay representation, such as the explanation of an HFEA member in 2002: “People know we’re not just a bunch of mad, profit-seeking doctors running around in white coats...The lay people are there to make sure the doctors don’t let science run ahead of what society wants.”³³²

This brief examination of the historical impulse behind lay membership helps us address the first set of questions around HFEA appointments: what ‘lay’ means in this context and what lay members are meant to bring to deliberations. ‘Lay’ in the HFEA context has been defined negatively, as the absence of medical or scientific *expertise* and a corresponding independence from the particular *interests* of this class of experts.

What is striking, then, about the rise of the lay member within public bodies like the HFEA is that it ‘solves’ a relatively narrow set of problems. All that has been said for lay membership so far is that it may prevent medical and scientific interests from

³³¹ For a sociological account of the power granted to ‘professionals’ in the medical field in the UK up through the 1970s, see Mary Ann Elston, “The Politics of Professional Power: Medicine in a Changing Health Service,” in *The Sociology of the Health Service*, ed. Michael Bury, Michael Calnan, and Jonathan Gabe (London: Routledge, 2002). For a more holistic account of the move towards subjecting the medical profession to public regulation, see Brian Salter, “Who Rules? The New Politics of Medical Regulation,” *Social Science & Medicine* 52, no. 6 (2001).

³³² Brendan Koerner, “Embryo Police,” *Wired Magazine*, February 2002, <http://archive.wired.com/wired/archive/10.02/biotech.html>.

hijacking deliberations and discussions that are “concerned essentially...with the protection of the public interest.”³³³ To be sure, this reclamation of authority from an elite class of technical experts is an important step and was, as I have tried to suggest briefly here, an historic one in UK public regulation of science and medicine. However, given the very background of pluralism upon which the Warnock Committee based its recommendations – in the terms taken up in Chapter III, the range of reasonable disagreement about the permissibility of novel reproductive technologies – the personal perspective of any given lay member or set of lay members may reasonably be contested by members of the public who hold conflicting comprehensive doctrines. Lay members, simply by virtue of their laity, cannot easily ‘represent’ the public interest against medical and scientific interests, for their personal views, while not those of the medico-scientific elite, are still contested. While meant to bring public perspectives to bear on issues once left to doctors or scientists, lay members still need a sharper way of thinking about how those public perspectives should be brought to bear. They need a way, in other words, to think about “representing” or “protecting” the public interest beyond simply embodying the ‘lay perspective,’ for there is no single lay perspective or set of lay interests.³³⁴

Here the reciprocal model hopes to provide that heretofore missing method of thinking or reasoning: the criterion of reciprocity stipulates that lay members engage alongside (notably) similarly-motivated scientific and medical expert-members in seeking mutually justifiable or public reasons upon which to base policy decisions. This means that the way in which lay members and expert members are asked to reason within the reciprocal model about the normative basis of policy decisions is not, in fact,

³³³ “Report of the Committee of Inquiry into Human Fertilisation and Embryology,” 76.

³³⁴ For an argument, drawing on in-depth interviews with lay members, that the concept of the ‘ordinary’ or ‘average’ citizen is misleading, see P. Lehoux, G. Daudelin, and J. Abelson, “The Unbearable Lightness of Citizens within Public Deliberation Processes,” *Social Science & Medicine* 74, no. 12 (2012).

different. This understanding is important for two reasons. First, it properly recognizes that simply because doctors and scientists may have some unique and conflicting interests due to their research or vocation, they are not incapable of thinking sincerely about the broader public interest. Indeed, the Warnock Committee's stark contrast of expert interests and lay interests risks overstating the gulf between the two and missing the many interests and goals that they share around the promotion of health and well-being. Second, in addition to reducing the emphasis on divergent *interests* of lay and expert members, it suggests that the lay member's main role (at least in setting standards rather than meeting them) is not tied to some special *expertise* that she brings in virtue of her understanding of the 'lay interest.'³³⁵

This suggests that many supposed 'problems' identified in the literature on lay membership are not as central and worrisome as we might imagine. So, on the one hand, many have worried about the true 'disinterestedness' of any lay person who would serve on a committee like the HFEA, which takes up a lot of time and offers relatively little compensation.³³⁶ On the account I have set out here, we can accept that this is the case: we will not find lay members that represent either a purely impartial perspective or that hold and reflect the 'interests of the public,' but this is not necessary (nor sufficient) to reason reciprocally. On the other hand, many note that lay members, in learning the facts about the issue facing the committee and thus coming to understand the issue more like the experts than the general public, are dispossessed of the very virtue underpinning their

³³⁵ Some have noted how lay members' knowledge of the patient experience may help, for instance, in constructing a consent form that patients can understand and engage with. I take this to be a relatively unproblematic issue of lay involvement for the purpose of standard meting.

³³⁶ Evans and Plows, "Listening Without Prejudice?"; Moore, "Public Bioethics and Public Engagement"; Tim Tenbensel, "Virtual Special Issue Introduction: Public Participation in Health Policy in High Income Countries – A Review of Why, Who, What, Which, and Where?," *Social Science & Medicine* 71, no. 9 (2010).

place.³³⁷ But we want decisions to be more informed than the view of the average lay member of the public, so it would be counterproductive to try to ‘protect’ lay members from gaining scientific or medical knowledge about the technologies under deliberation.

However, if the method of reasoning reciprocally ‘solves’ these problems by suggesting that lay members are not particularly valuable on a knowledge/expertise account nor on an interests/disinterested account, why have them at all? Lay members, I propose, may reprise their role in checking, not the interests or expertise of experts, but their ways and patterns of thinking about problems. Lay members bring a diverse *perspective* that is not constrained by the parochial concepts or disciplinary habits of approaching questions that are attendant to expertise. To use an example from the development of bioethics as an interdisciplinary field – and the perceived value of interdisciplinary work – it is clear that different disciplines think about problems in different ways. If we get a sociologist, anthropologist, philosopher, lawyer, and doctor in a room, we will initially be confronted with a diversity of conceptions of what we’re talking about when we talk about, for instance, patient autonomy, and what sort of considerations we need to attend to if we are to promote it. Even if doctors or scientists come to the table willing to reason reciprocally, rather than simply bargain in favour of their professional interests, they may sometimes fail to break out of these insular ways of approaching problems and thus be unable to seek out or recognize as relevant certain considerations. Here, the lay member is valuable because of her lack of disciplinary blinkers, not a lack of interest or expertise. This example also suggests the value of more general diversity of backgrounds on a committee like the HFEA.

³³⁷ Jill Thompson et al., “Credibility and the ‘Professionalized’ Lay Expert: Reflections on the Dilemmas and Opportunities of Public Involvement in Health Research,” *Health* 16, no. 6 (2012); Hogg and Williamson, “Whose Interests Do Lay People Represent?”

Beyond this account of lay membership, the reciprocal model suggests a more nuanced take on the HFEA's main *restriction* on appointments, lay and professional alike: that no appointees can be members of pro-life organisations. In Moore's interviews with members of the HFEA, the former chair explained why this is so:

I think that you must subscribe to the acceptability of IVF and the acceptability of embryo research. I do not think that you could sit on the Authority and exercise the kind of decision-making that we have to do if you were fundamentally opposed to the activities that we regulate.³³⁸

Membership in a pro-life organisation appears to serve as a heuristic for those who unjustifiably seek to impose their comprehensive moral views on others. However, the mere belief in the personhood of the embryo and a resulting motivation to prevent its destruction does not *immediately* rule one out as a person capable of reasoning reciprocally, of recognizing that others do not share this belief and thus accepting as properly 'relevant' or 'public' their arguments about the potential of embryo research to lead to life-saving treatments or the opportunity for the previously infertile to reproduce. One should be able to place great value on the embryo and yet still see countervailing, reasonable arguments. Indeed, those who are pro-life and yet support the legality of abortion when the mother's life or health is at stake at least implicitly recognize the existence and force of countervailing, public reasons that permit what they perceive as the active killing of persons.

Because of this, the reciprocal model would not advocate a total ban on those who identify as pro-life. Rather, membership should be restricted based on the acceptance of reasonable pluralism and the willingness to reason on the basis of reciprocity, questions that will need to be brought up in interviewing potential members. This is, to be sure, a no less restrictive requirement than one excluding pro-life members;

³³⁸ Moore, "Public Bioethics and Public Engagement," 205.

indeed, it is likely to rule out many who identify as secular and, for instance, seek to impose their fundamental beliefs in reproductive autonomy or the ultimate goal of maximizing welfare regardless of what values others hold. But it places these restrictions in the right way: it is less substantively restrictive, requiring only a willingness to pursue mutually justifiable reasons, whereas the pro-life restriction removes the possibility of certain comprehensive doctrines from entering the deliberative space at all. In fact, having such a pro-life member might help to identify mutually acceptable reasons, as well as to fully articulate deliberative disagreement when it persists. It would also go some way to preventing gerrymandering concerns; as it stands, pro-life citizens may understandably remain unconvinced that the HFEA attempts to locate reasons that all can agree are ‘relevant,’ since their own perspective is ruled out from the start. Indeed, Jonathan Montgomery, a prominent figure in UK bioethics and former chair of a number of arm’s length bodies including the Human Genetics Commission, has noted as much:

There is good reason to think that the way in which bioethical issues are discussed in public bodies and examined in the courts struggles to comprehend theological concerns...[D]ebates around the regulation of end of life care and assisted conception are cast in terms of whether there is evidence of harm that justifies constraining personal choice, not in terms of the intrinsic value of life...From the perspective of a public sector chairman, I am concerned where the ground rules of debate privilege some voices and create barriers to others being heard.³³⁹

The point is not that such views should be immune from scrutiny, or that they will always or even often be capable of withstanding the test of reciprocal reasoning, but that they should not be excluded from the outset.³⁴⁰

³³⁹ Jonathan Montgomery, “Public Ethics and Faith,” *Theology* 117, no. 5 (2014): 344-345.

³⁴⁰ Montgomery himself thinks it “appropriate to exclude from membership of the Human Fertilisation and Embryology Authority those who are so fundamentally opposed to research involving human embryos that they could not carry its licensing and supervisory roles” (346). But, again, I think those unwilling to carry out the justly deliberated decisions and roles of the HFEA will be excluded by reference to reciprocity, not their religious views.

Let us move on now to the question of the role, if any, that philosophers may justifiably play on arm's length bodies like the HFEA. As the above discussion of the origin and continued conception of lay representation on the HFEA should make clear, philosophers are lay members in the sense of not being scientific or medical experts, those who were traditionally involved in what Moore calls the 'technocratic model' of biomedical regulation, in which those who know the most about the facts of the field also set the normative rules.³⁴¹ But, as Archard notes, they nevertheless seem to be asked to sit on bodies like the HFEA because of some form of expertise that they bring to deliberations; in other words, to the extent that we want philosophers or ethicists to sit on such bodies, we must think that they have, and want them to provide, something more than the basic lay perspective that, say, an accountant might bring as a lay member.

One quality of expertise is that it is exclusionary: if we recognize that philosophers are 'experts' in the context of the HFEA, we also recognize that they either perform something better or know something more than other members. This troubles many to the extent that the philosopher's claim to expertise is understood as knowing better *what we ought to do*. The worry seems to be that recognizing the philosopher's expertise means that we should treat her moral judgements the way we treat the scientist's judgements: that we should treat them as authoritative. For instance, Moore worries that domination by 'ethics experts' on bodies like the HFEA is simply replacing the old model of domination by technical experts:

The central new problem that public bioethics raises for deliberative theory concerns the status and role of ethics expertise...By assuming the authority to...fram[e] ethical questions such that some kinds of concerns appear legitimately ethical while others are merely political or transient matters of public concern, public bioethics may be reproducing problems

³⁴¹ Alfred Moore, "Questioning Deference: Expert Authority in a Deliberative System" (Annual Meeting of the Midwest Political Science Association, Chicago, 2011).

of expert domination at the more rarefied level of the construction of ethical positions.³⁴²

This suggests that the expertise that philosophers bring to arm's length bodies is knowledge of the 'right' way of framing moral questions and knowledge of the 'right' kinds of ethical claims. In turn, recognizing this expertise, other members of arm's length bodies – be they technical experts or lay members – must defer to their judgements in these matters the same way the philosopher might defer to the scientist's judgment about scientific questions.

This understanding of the philosopher's expertise should sound familiar from the previous section's discussion of the critical model of ethics committees and the idea of epistocracy. In rejecting that model, that arm's length bodies should seek to base coercive political decisions on a comprehensive account of the truth, I rejected the idea that we ought to be ruled by the judgments of 'moral experts.' In a process of reciprocal reasoning, no moral expert can "silence"³⁴³ the non-expert on normative questions. To the extent, then, that the only expertise that the philosopher has to offer is a form of 'moral expertise' (regardless of whether or not you think philosophers are, in fact, moral experts), it is of very limited usefulness in the reciprocal model of ethics committees.

However, there does seem to be another type of expertise that philosophers possess that is valuable and valid in reciprocal deliberation. Specifically, although she may not claim to know better than others which reasons are public or mutually justifiable, her skill set may make her a better steward or facilitator of reciprocal reasoning than the average member. The largely non-controversial skills of the professional philosopher suggested in the previous section – a developed attention to

³⁴² Alfred Moore, "Public Bioethics and Deliberative Democracy," *Political Studies* 58, no. 4 (2010): 727.

³⁴³ I draw the conception of moral expertise as the ability to 'silence' the non-expert from Christopher Cowley, "Expertise, Wisdom and Moral Philosophers: A Response to Gesang," *Bioethics* 26, no. 6 (2012): 337.

consistency, bias, and invalid inference, past experience with the weighing of multiple, relevant values or considerations, and developed consideration of how empirical evidence relates to normative argument, are all skills that will help guide a process of deliberation. Indeed, given that the concepts of reciprocity, public reason, and deliberative disagreement are neither obvious nor simple to put into practice, a philosopher who could explain the meaning and implications of these ideas and guide deliberation back to reciprocity when it strays would be a valuable asset.

This supports an expert role for the professional philosopher as a steward of deliberation. By suggesting this, I do not mean to propose that a philosopher should always be the chair of such an arm's length body; in fact, in recognition of her basic equality with other members in the realm of judging the justifiability of various reasons and arguments, it may be more desirable that she be an ordinary member rather than chair. But her expertise in describing the framework of accountability and guiding deliberation should nevertheless be recognized and exercised. Furthermore, she may be well placed to draft the explanation of deliberative proceedings and grounds of decisions that I suggest in Chapter III as a vital component of accountability, both in providing a transparent account of decisions and in publicly providing the information that citizens require if they are to engage with and contest these decisions.

Having offered accounts of public consultation, lay membership, and philosophical expertise within the context of the HFEA that follow from the application of my public accountability framework, I turn now to considering and critiquing using the reciprocal model the way in which a particular decision was carried out.

3. The hard case of non-medical sex selection: what went wrong?

In this section, I examine the HFEA's policy reviews of sex selection, focusing on its most recent decision-making on this question. Using the reciprocal model of public ethics committees as a critical lens, I suggest some problems with the way in which the process of decision-making was carried out and how it may be improved. I have chosen the case of sex selection for two main reasons, both of which may best be explained by comparison of the policy reviews of sex selection with the more recent policy review of mitochondrial donation. First, in the case of mitochondrial donation, the HFEA's role was to provide advice to Parliament rather than make an executive decision itself. The *Human Fertilisation and Embryology Act 1990* (as amended in 2008) only allows eggs and embryos that have not had their nuclear or mitochondrial DNA altered to be used in clinical treatment. As such, allowing mitochondrial donation required a change in the law which, as the HFEA's report to the Government noted, made the decision "quite properly, one which only Parliament can take."³⁴⁴ In contrast, the HFEA's last policy review of sex selection, in 2002-2003, fell within the arm's length body's regulatory domain, and thus the outcome of the review was a politically authoritative decision rather than advice to government.³⁴⁵

Second, although there was a vocal minority who raised fundamental ethical objections to creating babies with "three [genetic] parents" or altering the germ line, the HFEA found using a representative survey of the public and deliberative public workshops that "there is general support for permitting mitochondria replacement in the

³⁴⁴ Human Fertilisation and Embryology Authority, "Mitochondria Replacement Consultation: Advice to Government," March 2013, http://www.hfea.gov.uk/docs/Mitochondria_replacement_consultation_-_advice_for_Government.pdf.

³⁴⁵ However, the HFEA's decision to ban non-medical sex selection was ultimately enshrined in statute by Parliament during the 2008 amendments to the HFE Act.

UK.”³⁴⁶ Most ethical concerns were about the safety of the technique; so long as additional planned research showed the risks to be minimal and so long as regulations were put in place to manage any remaining uncertainties about safety or unexpected consequences, the public did not overwhelmingly register moral objections. In contrast, in the case of non-medical sex selection, the majority of the public *opposed* the new technology, not on the basis of safety that might be assuaged by further research and evidence, but on the basis of far more morally contested values and arguments about the proper valuation of children and the ‘harm’ that being selected on the basis of sex might cause. As my analysis of the case below will suggest, these arguments were not necessarily ones that members of the HFEA found convincing. As such, they seemed to be left in a difficult position of having to decide ‘against’ either public consultation results or critical reasoning.

Both of these differences between the mitochondrial donation and sex selection cases illustrate why the latter represents a particularly ‘hard’ decision for the HFEA. First, the HFEA itself, rather than a superior and democratically accountable body, had to make the final, politically coercive decision. And second, there seemed to be a gulf between the reasons identified as legitimate by the public and those identified by the HFEA after careful consideration. So while it may not be a representative or even a particularly common problem – a full analysis of HFEA decisions would be required to determine this – it does represent the type of tension between philosophical reasoning and public representation that I am interested in working through.

In the 1980s, largely due to the work of Derek Parfit, the ‘non-identity problem’ became widely recognized amongst academic philosophers as a troubling puzzle for our

³⁴⁶ Human Fertilisation and Embryology Authority, “Mitochondria Replacement Consultation: Advice to Government,” 4.

intuitions about harm to future persons.³⁴⁷ The problem has particularly strong implications for reproductive choices. Consider the oft-cited case of a deaf couple that desires to have a deaf child.³⁴⁸ To assure this result, the couple uses IVF and pre-implantation genetic diagnosis (PGD), creating two viable embryos. Embryo A will result in a deaf but otherwise healthy child. Embryo B will result in a healthy, hearing child. Should the couple be permitted to implant Embryo A rather than Embryo B? Although the deaf community cites the vibrancy of deaf culture, many people's primary intuition is that selecting Embryo A would be a 'bad' choice, that there is something wrong with choosing a disabled child. But, as the non-identity problem suggests, this choice cannot be understood as making Embryo A 'worse off.' For if the couple makes any other choice – to implant Embryo B or not have a child after all – then Embryo A will never come into existence as a person. More generally, unless a condition selected for using PGD is so terrible as to create a person with a life not worth living (which deafness is clearly not), reproductive decisions about which persons to bring into existence cannot 'harm' the resulting persons. And as such, a commitment to protecting the 'welfare of the child' cannot justify prohibitions on creating children with disabling conditions, so long as they do not end up with a life not worth living.

As it often takes some time for academic ideas to reach the policy setting, this philosophical development did not immediately show impact in the early work of the HFEA. In 1993, describing in a public consultation document why it allows sex selection for medical reasons, the HFEA specifically notes, "A person with a disease will often come to accept it but may rather have been born healthy," failing to recognize that had

³⁴⁷ Derek Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1984): 351-378.

³⁴⁸ Savulescu, "Deaf Lesbians, 'Designer Disability,' and the Future of Medicine"; John A. Robertson, "Extending Preimplantation Genetic Diagnosis: The Ethical Debate Ethical Issues in New Uses of Preimplantation Genetic Diagnosis," *Human Reproduction* 18, no. 3 (2003); Neil Levy, "Deafness, Culture, and Choice," *Journal of Medical Ethics* 28, no. 5 (2002).

sex selection been used to prevent the birth of a child with a congenital disorder, an entirely different person would have been born – not simply the same hypothetical person less the disease.³⁴⁹ On the flip side, in explaining the arguments against sex selection for non-medical reasons, the HFEA notes, “Should a child not of the selected sex be born, then its parents may treat it less well than if it had been of the selected sex. The welfare of that child could be affected,” expressing a concern arising from the questionable accuracy of some sex selective techniques in the early 1990s.³⁵⁰ Setting aside the question of whether parents do in fact treat their children less well if they are not of the desired sex, this reasoning overlooks the fact that the decision to use sex selection cannot strictly harm the resultant child, for he is brought into existence by virtue of that choice. While merely one among several reasons, concern for the welfare of the child underpinned the HFEA’s decision in 1993 to ban sex selection for non-medical, or ‘social,’ reasons.

A decade later, the HFEA revisited the issue of sex selection for social reasons. In contrast to the HFEA’s 1993 policy review, the 2002-2003 public consultation document exhibited a clear understanding of the non-identity problem, noting:

When one thinks of sex selection one has to be careful not to commit the fallacy of thinking that by using these techniques one is choosing the sex of ‘an embryo’ or ‘a child’. The techniques do not affect the sex of embryos or children who already exist: the choice is either about which embryos will be created in the first place (sperm sorting) or about which embryos, from among those that already exist will be transferred (PGD).³⁵¹

³⁴⁹ Human Fertilisation and Embryology Authority, “Sex Selection: Public Consultation Document,” January 1993, http://www.hfea.gov.uk/docs/Sex_Selection_Consultation_1993.pdf, 7.

³⁵⁰ *Ibid.*, 8.

³⁵¹ Human Fertilisation and Embryology Authority, “Sex Selection: Choice and Responsibility in Human Reproduction,” January 2003, http://www.hfea.gov.uk/docs/Sex_Selection_choice_and_responsibility.pdf, 9.

During the same period, the Ethics and Law Advisory Committee (ELAC)³⁵² – a former sub-committee of the HFEA composed of a smaller set of its regular members – embarked upon a far more in-depth analysis of the non-identity problem’s implications for the ethics of sex selection. In contrast to the HFEA’s 1993 justification for the use of sex selection for medical reasons, which was based on the welfare of the resulting child, the ELAC reasoned:

Since PGD is not a form of gene-therapy it cannot be straightforwardly represented as a practice which produces a healthy child who would otherwise have been likely to be seriously disabled and thus as a practice warranted by the requirement to attend to the welfare of the child...Hence it is the interest of the parents in having a healthy child...that warrants the practice.³⁵³

Here the ELAC resolved that only a consideration of the interests of the *parents*, not the child, can justify the use of PGD to select against disease and disability, judging that this parental interest in having a healthy child outweighs considerations about the ‘special status of the embryo’ that might otherwise suggest the wrongness of destroying certain embryos based on their genetic makeup.

Delving deeper, the ELAC engaged with Parfit’s distinction between person-affecting and non-person-affecting (or impersonal) principles, reasoning that “dysgenic” practices like selecting a deaf child cannot be rebutted using arguments about the welfare of the child. Rather, “one has to argue that this *outcome* is worse than that in which a child with hearing is born, and that this fact legitimates regulatory intervention to prevent it.”³⁵⁴ It is true, in other words, that we cannot reason from a claim about the welfare of a child to a prohibition on selecting a deaf child, as no person-affecting harm

³⁵² See Chapter II for a discussion of this subcommittee and the role of subcommittees more generally in the HFEA.

³⁵³ HFEA Ethics and Law Committee, “Ethical Values and the HFEA,” December 2003, http://www.hfea.gov.uk/docs/ELC_2_Dec03.pdf, 7-8.

³⁵⁴ *Ibid.*, 8.

is perpetrated. Yet other arguments, for instance, certain consequentialist arguments that rely on an impersonal principle that we should maximize aggregate welfare, may nevertheless allow us to reason that it is *better* to bring a hearing child into existence than a deaf one. In sum, the evidence from these reports confirms the HFEA's familiarity with and apparent acceptance by at least some members of the non-identity problem's implications for its regulatory work.

Yet despite this clear understanding of the complications, the weight afforded to the welfare of the child concern was if anything more pronounced in the HFEA's second review of sex selection. The final report, released in 2003, justified the decision to maintain the ban on sex selection for social reasons by citing the welfare of children born using sex selective techniques and the overwhelming public endorsement of this line of reasoning:

In reaching a decision we have been particularly influenced by the considerations set out above relating to the possible effects of sex selection for non-medical reasons on the welfare of children born as a result, and by the quantitative strength of views from the representative sample polled by MORI and the force of opinions expressed by respondents to our consultation.³⁵⁵

Arguments about the wrongness of sex selection based on 1) interference with divine will or the course of nature, 2) the perpetuation of sex discrimination or sexual stereotypes, and 3) the 'slippery slope' to selecting embryos based on personality, intelligence, or appearance were also raised in this review, but given less weight in the HFEA's justification than the welfare effects on the resulting children. Seemingly incongruous in the wake of both the ELAC's and the larger HFEA's development of thought on the matter, the "considerations set out above" that the HFEA refers to in this

³⁵⁵ Human Fertilisation and Embryology Authority, "Sex Selection: Options for Regulation - A Report on the HFEA's 2002-03 Review of Sex Selection Including a Discussion of Legislative and Regulatory Options," 2003, http://www.hfea.gov.uk/docs/Final_sex_selection_main_report.pdf, 36.

justification are precisely the type that the non-identity problem would cast aside. They are worries drawn from the public consultation that, for instance, “parents would foist their gender expectations on the child,” and that the level of “both physical and psychological” risk to the child from such potential consequences is too high.³⁵⁶

On the one hand, expert members of the HFEA – particularly those that sit on the ELAC – now appear not only familiar with the non-identity problem but convinced by its challenge to evaluating the welfare of future persons as a method for determining whether their creation is permissible. Yet on the other hand, the HFEA’s review clearly found, through public consultation, focus groups, and a poll of a representative sample of the British public, that the public understand the issue in this way. The HFEA was thus confronted with the knowledge that 69% of the population opposed sex selection for non-medical reasons, but that many did so on the basis of reasoning that it found to be flawed.

The apparent incongruity between the HEFA’s careful reasoning about future persons and its final rejection of sex selection for social reasons on the grounds of the welfare of the child suggests a practical uncertainty about the role that such ‘expert’ philosophical reasoning should play in public policymaking. John Harris draws attention to the way in which the HFEA appears to ‘flip-flop’ on its principled commitments over the course of this second review of sex selection. He notes how the HFEA begins with a commitment to the ‘liberal presumption’ in its report³⁵⁷: “[Reproduction] is an area of private life in which people are generally best left to make their own choices and in which the State should intervene only to prevent the occurrence of serious harms, and

³⁵⁶ Ibid., 16-22.

³⁵⁷ He refers to it as both the democratic and the liberal presumption, but the presumption is most clearly aligned with liberal principles in its endorsement of the harm principle to the end of creating a sphere of reproductive autonomy. Given the distinction between aggregative and deliberative democratic theories, the label ‘democratic presumption’ is perhaps unclear. Harris, “Sex Selection and Regulated Hatred.”

only where this intervention is non-intrusive and likely to be effective.”³⁵⁸ Yet the HFEA’s justification for prohibiting sex selection, only pages later in the report, states:

[T]here is very *wide-spread hostility* to the use of sex selection for non-medical reasons. By itself this finding is not decisive; the fact that a proposed policy is widely held to be unacceptable does not show that it is wrong. But there would need to be substantial demonstrable benefits of such a policy if the state were to *challenge the public consensus* on this issue. In our view the likely benefits of permitting sex selection for non-medical reasons in the UK are at best debatable and certainly not great enough to sustain a policy to which the great majority of the public are strongly opposed (emphases added).³⁵⁹

The burden of proof, in other words, is shifted in this statement from those who would prohibit sex selection to those who would use it. It also seems to commit the HFEA to a deeply illiberal principle for future policy reviews. It sets up a test in which there must be a ‘substantial demonstrable benefit’ of a policy in order for it to prevail over ‘the public consensus.’ Because of this unexplained reversal of its liberal presumption, Harris thus draws the conclusion that it represents the HFEA’s “attempt to formalize the tyranny of the majority.”³⁶⁰

Harris ultimately claims that ‘what went wrong’ in this case is that the HFEA failed to base its decision on the right, *substantive* consideration, reproductive autonomy, given his commitment to the ‘liberal presumption’ of the harm principle that the law should only intervene in the decisions of individuals if they cause demonstrable harm to others. In contrast, I would like to examine how the HFEA got the method or *process* of reasoning wrong in this instance, how it failed to reason reciprocally.

³⁵⁸ Human Fertilisation and Embryology Authority, “Sex Selection: Options for Regulation - A Report on the HFEA’s 2002-03 Review of Sex Selection Including a Discussion of Legislative and Regulatory Options,” 33.

³⁵⁹ *Ibid.*, 36.

³⁶⁰ Harris, “Sex Selection and Regulated Hatred,” 294.

The primary barrier to reciprocal reasoning in this decision was the improper conception and, as a result, design of public involvement processes. In particular, the justification of the decision on the basis of “the quantitative strength of views from the representative sample polled by MORI and the force of opinions expressed by respondents” represents a collapse of public accountability in arm’s length governance, perhaps an ironic one given the clear desire of the HFEA to demonstrate consideration of the public *by virtue* of this justification. Reference to a majority or even a supermajority in popular opinion can not show on its own that a position is mutually justifiable in the way demanded by reciprocity, and no magnitude or ‘force’ with which a person or people hold an opinion can serve to make it public and justificatory in the face of reasonable pluralism if it is particular to one of many, conflicting conceptions of the good life. Furthermore, as discussed in my rejection of the mirror model of ethics committees, reliance on such quantitative polling results conflicts not only with the goal of reciprocity, but also with the fundamental purpose of setting these decisions at arm’s length: considered, *non-majoritarian* judgements of issues surrounding reproductive technologies and research. To rely upon such majoritarian justifications in this context is thus not only non-reciprocal, but also inconsistent when considered against the HFEA’s existential justifications.

In contrast to this representative opinion poll, the HFEA also commissioned a set of deliberative focus groups, which were “drawn from a wide range of ages and social backgrounds in hour-long discussions, four of the groups reconvening after a week for a further session. Additionally four specific ethnic minority groups were held.”³⁶¹ Such public participation methods are more well-suited to the goal of eliciting the range of

³⁶¹ Human Fertilisation and Embryology Authority, “Sex Selection: Options for Regulation - A Report on the HFEA’s 2002-03 Review of Sex Selection Including a Discussion of Legislative and Regulatory Options,” 12.

public values on an issue, given the deliberative space in which participants can learn about the technique, why others are opposed to or support it, and thereby test and refine their views on the relevant considerations. In fact, although less weighty in the HFEA's reasoning than the representative poll, these deliberative groups revealed a much different picture of the public perspective on sex selection, a truly broader set of considerations that members of the public recognize as relevant and compelling in decision-making about whether to legally ban or permit sex selection.

The report summarising the findings of the deliberative focus groups suggests that members of the public recognize two separate sets of issues: "issues related to being a citizen and/or consumer in our society" and "issues that were more concerned with moral or religious arguments and feelings."³⁶² Participants struggled with the idea of justifying restrictions on sex selection to others as *fellow citizens*, struggled – in the terms I have been using – to give reasonable justifications for banning sex selection:

Whatever their 'end' position, the impact of couching the debate in terms of citizenship, and thinking about the issue from this standpoint meant that respondents felt they had to produce very robust arguments against sex selection for social reasons. They felt the onus was on them to justify their seeming intolerance, and they struggled to do so. However, at the end of the day, most felt that the key decision-making issues were not those of citizenship, but moral or to do with their definition of a family and parenthood."³⁶³

Thus, although the majority in the focus groups were still ultimately opposed to sex selection for non-medical reasons, this method of involvement illustrates the far more complex set of values and considerations that citizens recognize as relevant to such discussions. In particular, the report notes how even those who found other moral

³⁶² Counterpoint UK, "Sex Selection - Policy and Regulatory Review: A Report on the Key Findings from a Qualitative Research Study Prepared for the Human Fertilisation and Embryology Authority" (London, October 2002), [http://www.hfea.gov.uk/docs/Appendix_E_-_Qualitative_Research_Findings_report_of_research_conducted_by_Counterpoint_\(UK\)_Ltd.pdf](http://www.hfea.gov.uk/docs/Appendix_E_-_Qualitative_Research_Findings_report_of_research_conducted_by_Counterpoint_(UK)_Ltd.pdf), 14.

³⁶³ *Ibid.*, 22.

considerations, based on their definition of family, to be overriding, recognized the force or weight of countervailing considerations, like reproductive autonomy:

Most could also see the argument for why some parents would want to select the sex of their child for non-medical reasons...parents in particular felt that it was very difficult to argue the case against giving one couple their heart's desire when no-one was harmed – however uncomfortable it made them feel personally.³⁶⁴

This is a picture of public values and public perspectives that is not adequately captured in the method of representative polling. It is also the type of picture that the reciprocal model needs to capture from public consultation; members of the HFEA, as I shall discuss shortly, must be able to use these results to search for mutually acceptable justifications for policies. Knowing simply, based on revealed preferences in a poll, what individuals find to be the 'weightiest' consideration sheds light only on a limited section of their conflicting, comprehensive doctrines rather than illuminating potentially shared portions.

The fact of reasonable pluralism, however, means that we are unreasonable when we try to base public, coercive decisions on our own comprehensive accounts of the good life. Because of this, it is the hard task of the HFEA to take on that 'citizenship perspective' that participants struggled with. The personal discomfort that focus group participants registered in taking this perspective illustrates that reciprocal reasoning is truly a demanding process. Indeed, although participants "hoped that the regulators could come up with good reasons why it should be banned within this [citizenship] perspective," they supported delegating this decision to the arm's length context, "declar[ing] that the ideal would be for such regulation to be as apolitical as

³⁶⁴ Ibid., 27.

possible...involv[ing] a ‘panel’ of experts and ‘independent’ people, including the public, or regularly consulting the public.”³⁶⁵

How did this arm’s length decision-making process fare after the return of public consultation results, taking as given that the reliance on the inferior method of polling rather than focus groups in itself hinders reciprocal reasoning? First, we may ask how deliberators should have exercised their ‘expert’ understanding of the non-identity problem in relation to claims about the harm to sex-selected children. I take it that the challenge posed by the non-identity problem is a *logical* one: it is based on standards of valid inference that are mutually acceptable. It is worth noting in relation to this that, upon consideration, the majority of the participants in the focus groups understood non-medical sex selection as *harming* no one; indeed, participants moved quickly past questions of harms to consider values of autonomy, family, and parenthood. Because of this logical challenge, it would have been in line with reciprocal reasoning to ultimately reject as unreasonable the claim that a parent’s decision to bring a female embryo into existence rather than a male embryo harms their future child. Of course, in line with publicity and fostering deliberative resistance, this reasoning would have to be carefully explained in a publicly accessible way.

Moving past this discrepancy between the expert deliberators and the public focuses deliberation on the far harder but more important question of whether reciprocal reasoning can locate a determinate answer based on the range of values around family, parenthood, and reproductive autonomy that were identified in public consultation. One potential answer, supporting the liberal presumption, is that the only mutually acceptable reason here, once we recognize the lack of harm, is that of reproductive autonomy: given

³⁶⁵ Ibid., 24. Most participants had not heard of the HFEA and did not know that such a regulatory body was already in effect.

that we have reasonable but conflicting accounts of how parents ought to relate to their children, none of *my* arguments about the definition of family can serve to justify limits on *your* reproductive decisions. While I am inclined to see the issue in this way, my goal here is not primarily to support or reject particular outcomes of reciprocal reasoning, but rather flesh out what the process should look like.

So, for instance, opponents of sex selection might note how, even if we have conflicting conceptions of the family, we all recognize the importance of the family as an institution within society, that it is fundamental, in Rawls' terms, to the "ordered reproduction of political society over time."³⁶⁶ We rely heavily on the family to impart our political values of freedom and equality of persons, which provide the very basis for an overlapping consensus on public justification. Arguments in favour of banning sex selection based on its potential, for instance, to promote sexual discrimination, particularly within cultural groups that place a priority on having male children, and thus threaten this ordered reproduction of political society are therefore based in *public* reasons as much as arguments about reproductive autonomy. This is the type of argument that one might give against sex selection in attempting to reason reciprocally. Say that it convinced the HFEA that there were arguments that were properly reciprocal both for and against sex selection, about autonomy and the family. What then? Any policy position would ultimately require weighing more heavily one of these two reasons, in a way that the HFEA may not be capable of justifying on mutually acceptable grounds: this is the problem, discussed in Chapter III, of the *inconclusiveness* of public reasoning.

My framework suggests that, procedurally, the next step for the HFEA in this position would be an attempt at moral accommodation. What might this look like here?

³⁶⁶ Rawls, *Political Liberalism*, 243.

Even if both sides of the argument recognize their own weighing of the values of autonomy and family to be the most reasonable one, they might be able to come to a moral compromise in spite of this deliberative disagreement. For example, they may all recognize that sex selection for family balancing purposes – to have a daughter after having one or more sons, say – is a particularly compelling argument for reproductive autonomy, given that raising boys and girls are different experiences for a parent, each of which he may value greatly. They may also recognize that this rationale for sex selection will not, or at least is far less likely, to promote sexually discriminatory values or other social values that would undermine the stability of political society. So the HFEA may be able to come to a moral compromise in permitting sex selection for family balancing but banning it in cases where parents simply want a son or a daughter for their first child.

If agreement could not be reached on a compromise position, it would be acceptable on my framework to reach a decision by voting, in which HFEA members base their votes on the arguments they find most convincing after this process of reciprocal reasoning. Their publicly accessible explanation of the decision-making process should outline the process as well as the two options in the vote (or three, if the family balancing option is included) and the reasons underpinning both the majority and minority positions. In addition, even if an HFEA ‘appeals process’ is not a practically or financially feasible option, the document should suggest when the policy will be considered for review (perhaps in five years time). The clarity of the document and this option for future challenge provide the means by which those who dispute the decision may engage with the arguments and contest them in due time, which meets the condition of fostering deliberative resistance in my public accountability framework.

Although at times a very crude overview of what are quite sophisticated and difficult issues, I hope this consideration of the case of sex selection has given some body to my conditions of public accountability: how they may shape the processes of public involvement and deliberative decision-making within the context of a specific determination. This fleshes out the previous work of the chapter, which provided the proper justifications based on a reciprocal model of ethics committees for public consultation as well as the appointment of both lay members and philosophers. I conclude this chapter having given a fairly holistic application of my framework of public accountability while also resolving the practical tensions from Chapter II around the HFEA's 'method' of reasoning and the uncertainty about its appropriate relationship to the public.

Chapter V. The Role of Patients in NICE: Providing and Contesting Evidence

Because the HFEA lacked a transparent and legitimate method of reasoning for the arm's length context, it provided an ideal setting in which to flesh out the heart of my framework of public accountability: the first condition of reasoning reciprocally. In addition to showing how reciprocity can guide the reasoning of a deliberative committee, the HFEA case study provided the opportunity to show how the theoretical claims from Chapter III about reasonable pluralism, public reason, and deliberative disagreement (upon which my framework is based) entail specific, justified conceptions and roles for public consultation, lay membership, and philosophical expertise, usefully working through these often ambiguous concepts. NICE presents a substantially different context, in that it has already explicitly committed to accountability for reasonableness as its method of reasoning about resource allocation.³⁶⁷ Indeed, many cite the Institute as a prime example of how Daniels' and Sabin's four conditions of publicity, relevance, challenge/revision, and enforcement may be adopted in a practical policy making setting.³⁶⁸ Even those who remain critical of NICE's ability to meet these conditions

³⁶⁷ National Institute for Health and Care Excellence, *Social Value Judgements: Principles for the Development of NICE Guidance*.

³⁶⁸ See Anthony J. Culyer, "NICE's Use of Cost Effectiveness as an Exemplar of a Deliberative Process," *Health Economics, Policy, and Law* 1, no. 3 (2006); Norman Daniels and James E. Sabin, "Accountability for Reasonableness: An Update," *BMJ* 337 (2008); Suzanne Hill et al., "Technology Appraisal Programme of the National Institute for Clinical Excellence: A Review by WHO" (Copenhagen: WHO Regional Office for Europe, 2003), <http://apps.who.int/medicinedocs/documents/s19194en/s19194en.pdf>; Craig R. Mitton et al., "Centralized Drug Review Processes: Are They Fair?," *Social Science & Medicine* 63, no. 1 (2006).

nevertheless confirm, in virtue of such criticism, the Institute's central goal of being accountable for its reasonableness.³⁶⁹

The influence of NICE's commitment to accountability for reasonableness is particularly apparent in its technology appraisal role, the function by which it regulates public access to emerging biotechnologies and the one with which I am thus concerned. Technology Appraisal Committees are instructed to use the substantive values set out in NICE's guiding document *Social Value Judgements* in reaching a decision, which in turn draws upon reports of the deliberative Citizens Council, a group of 30 'ordinary' and demographically representative members of the public.³⁷⁰ Given that Technology Appraisal Committee members are usually selected for their scientific expertise, this role of the Citizens Council appropriately recognizes that "[e]xpert scientists and clinicians...have no special legitimacy to impose their own social values on the NHS. Those should reflect those held by people who are using, or who will use, the service."³⁷¹ Nevertheless, the NICE Board must first explicitly consider and endorse these public values before they are incorporated into *Social Value Judgements*, providing a check against simplistic and problematic 'mirroring' of public preferences or values that have not been carefully scrutinized, a key concern in the HFEA case. Thus the Council embodies a properly constrained form of public consultation, albeit a different form from that recommended in the HFEA context: one which functions to elicit and deliberate on a range of public values around healthcare but that ultimately can only feed into the

³⁶⁹ Michael Schlander, "The Use of Cost-Effectiveness by the National Institute for Health and Clinical Excellence (NICE): No(t Yet An) Exemplar of a Deliberative Process," *Journal of Medical Ethics* 34, no. 7 (2008); Keith Syrett, "Deconstructing Deliberation in the Appraisal of Medical Technologies: NICEly Does It?," *The Modern Law Review* 69, no. 6 (2006).

³⁷⁰ Michael Rawlins, "Background to NICE's Citizens Council," in *Patients, the Public and Priorities in Healthcare*, ed. Peter Littlejohns and Michael Rawlins (Oxford: Radcliffe Publishing, 2009).

³⁷¹ Michael D. Rawlins, "Pharmacopolitics and Deliberative Democracy," *Clinical Medicine* 5, no. 5 (2005): 472.

reciprocal reasoning of NICE in its attempt to base decisions on reasons that all can accept are ‘relevant.’³⁷²

Apart from using values drawn from the Citizens Council and enshrined in *Social Value Judgements*, a way of addressing accountability for reasonableness’s relevancy condition, Technology Appraisal Committees exhibit publicity in their openness about both the processes and grounds of their decisions. The former are delineated in the *Guide to the Process of Technology Appraisal*,³⁷³ which sets out the stages of a typical technology appraisal, and *Guide to the Methods of Technology Appraisal*,³⁷⁴ which sets out the principles and technical methods for calculating clinical and cost effectiveness. The grounds of the decision to recommend or not recommend a specific technology are laid out in a publicly available and accessible document called a Final Appraisal Determination. And, attending to the challenge and revision condition, patient groups, primary care trusts, and pharmaceutical companies are afforded the opportunity to appeal this determination, by either disputing the substantive grounds of the decision or citing a failure of the Appraisal Committee to adhere to the processes set out publicly by NICE.

As I explained in Chapter III, the procedural conditions set out by Daniels and Sabin are founded on a substantive commitment to reciprocity, as evidenced by its second condition to seek out reasons that fair-minded people can accept are ‘relevant’ to

³⁷² For analysis and general support of the Citizens Council as a way of justifiably involving the public within a deliberative democratic framework, see Julia Abelson et al., “Bringing ‘the Public’ into Health Technology Assessment and Coverage Policy Decisions: From Principles to Practice,” *Health Policy* 82, no. 1 (2007); Celia Davies et al., “Opening the Box: Evaluating the Citizens Council of NICE: Report Prepared for the National Co-ordinating Centre for Research Methodology, NHS Research and Development Programme,” 2005, http://irep.ntu.ac.uk/R/?func=dbin-jump-full&object_id=188851&local_base=GEN01. For Daniels’ own defense of the Citizens Council’s contribution to accountability for reasonableness, see Norman Daniels, “Accountability for Reasonableness and the Citizens Council,” in *Patients, the Public and Priorities in Healthcare*, ed. Peter Littlejohns and Michael Rawlins (London: Radcliffe Publishing, 2009).

³⁷³ National Institute for Health and Care Excellence, “Guide to the Processes of Technology Appraisal” (NICE, September 2014), <https://www.nice.org.uk/article/pmg19/>.

³⁷⁴ National Institute for Health and Care Excellence, “Guide to the Methods of Technology Appraisal 2013” (NICE, April 2013), <https://www.nice.org.uk/article/PMG9/>.

the decision at hand. In adopting an accountability for reasonableness framework, then, NICE has already accepted the kind of reciprocal reasoning that I advocate for use in the arm's length context, as well as procedural requirements that follow from reciprocity, like publicity and the availability of avenues for appeal and revision. As such, I broadly endorse NICE's approach to public accountability, and the work of this chapter is not geared towards radically disrupting that framework. Similarly, I do not wish to simply reiterate the demands of reciprocity for a method of reasoning or for public involvement and member appointments, as I did within the HFEA chapter.

Rather, I use the case study of NICE to tackle the practical challenges in arm's length governance that endure in spite, and sometimes *because*, of the implementation of a public accountability framework that is largely comparable to my own. Problems persist in the arm's length context even after committing to reciprocity, setting out a fair process that all can agree to, and identifying a set of guiding substantive values that all reasonable people can recognize as 'relevant,' steps that NICE has already taken in structuring its technology appraisal process. I focus on overcoming some of these residual challenges in this chapter. In the NICE context, as I illustrated in Chapter II, one persistent problem is the uncertainty about the role that patient groups and the 'patient voice' should play in technology appraisals, given the perceived tension between involving and accounting to patients in decisions that will profoundly affect them and attending to clinical data and economic models of cost-effectiveness.

It is important to note that in focusing on this problem and the decision-making of Technology Appraisal Committees rather than reiterating the work of the HFEA chapter, I take as provisionally settled the judgements made by the NICE Board, in concert with the Citizens Council, that are enshrined in *Social Value Judgments* and by which Appraisal Committees are bound. In particular, I do not take issue herein with

NICE's current method of evaluating health gains and determining when a technology is 'value for money.' NICE uses the widely recognized unit of quality adjusted life years (QALYs) as the method for valuing health gains, and the use of a 'cost per QALY threshold' as the determinant of whether an intervention is a cost-effective use of public money. Rigorous debate continues about 1) whether the use of the QALY is justified,³⁷⁵ and 2) even if it is, whether the current threshold is set in the right place.³⁷⁶ However, I take it that even those who object to QALYs or the current threshold agree that, in the face of limited healthcare resources, considerations about the 1) clinical effectiveness and 2) cost of treatment are *relevant* considerations in attempting to make resource allocation decisions. We can agree that it would be unreasonable, in other words, to insist on publicly funding a treatment for which we have no evidence of benefit, or without any regard to the way in which that decision would deplete the healthcare resources available to others who are also in need.

A third relevant consideration, which NICE endorses alongside clinical and cost effectiveness in *Social Value Judgements*, is that these two considerations are not the *only* determinants of whether funding a certain treatment is permissible: there are other values that may be countervailing. A common example of this consideration's relevancy is that we would rather improve quality of life for a small group of people with a severe illness than treat a massive group of people with mild headaches, even if it produces fewer QALYs.³⁷⁷ In other words, maximizing QALYs is not the *only* thing that we think matters in rationing healthcare. NICE currently considers that fair distribution – in

³⁷⁵ Joanna Coast, "Is Economic Evaluation in Touch with Society's Health Values?," *BMJ* 329, no. 7476 (2004); John Harris, "Nice and Not so Nice," *Journal of Medical Ethics* 31, no. 12 (2005); John Harris, "It's Not NICE to Discriminate," *Journal of Medical Ethics* 31, no. 7 (2005).

³⁷⁶ Karl Claxton et al., "Methods for the Estimation of the National Institute for Health and Care Excellence Cost-Effectiveness Threshold," *Health Technology Assessment* 19, no. 14 (2015); Karl Claxton et al., "Causes for Concern: Is Nice Failing to Uphold Its Responsibilities to All Nhs Patients?," *Health Economics* 24, no. 1 (2015).

³⁷⁷ Koonal K. Shah, "Severity of Illness and Priority Setting in Healthcare: A Review of the Literature," *Health Policy* 93, no. 2–3 (2009).

addition to clinical and cost-effectiveness – also requires avoiding discrimination on the basis of race or sex, attending to socio-economic health inequalities, promoting innovation in the pharmaceutical industry, and extending life at the end of life for small patient groups with incurable illnesses.³⁷⁸

Principle	Summary of NICE's explanation
1. Clinical effectiveness	Should not recommend an intervention if there is no evidence, or not enough evidence, on which to make a clear decision.
2. Cost effectiveness	Should take cost effectiveness into account.
3. Beyond effectiveness	Should consider factors beyond clinical and cost effectiveness, such as fair distribution of health resources across society.
4. QALY thresholds	Should explain why interventions with incremental cost effectiveness ratios (ICERs) below £20,000 per QALY gained are judged to be not cost effective; and why interventions with ICERs of more than £20,000 to £30,000 per QALY gained is cost effective.
5. Limits of patient choice	Should not allow conceptions of patient autonomy or individual choice to dictate the recommendation of interventions that are not effective or not cost effective.
6. Responsive but ultimately responsible	Should consider and respond to comments it receives but must ultimately exercise own judgement in recommendations.
7. Avoiding discrimination	Should only recommend that use of an intervention is restricted to a particular group of people within the population under narrow set of circumstances.
8. Premium on reducing health inequalities	Should actively consider reducing health inequalities including those associated with sex, age, race, disability and socioeconomic status.

Table 2. NICE's Eight Substantive Principles from *Social Value Judgements*, as summarized by author. National Institute for Health and Care Excellence, "Social Value Judgements: Principles for the Development of NICE Guidance (2nd Edition).

As I argued in Chapter III, it is likely that in resource allocation we will be able to agree on what broad reasons are relevant but may nevertheless – despite attempting to reason reciprocally – continue to disagree about how to order these mutually justifiable reasons. We can understand the principles in *Social Value Judgements* (see Table 2),

³⁷⁸ National Institute for Health and Care Excellence, "Appraising Life-Extending, End of Life Treatments," Supplementary Advice to Appraisal Committees, (July 2009); Christopher McCabe, Karl Claxton, and Anthony J. Culyer, "The NICE Cost-Effectiveness Threshold: What It Is and What That Means," *Pharmacoeconomics* 26, no. 9 (2008).

particularly the first three on clinical effectiveness, cost effectiveness, and the existence of *some* non-effectiveness considerations, as setting out those broad values or considerations recognized as relevant to resource allocation decisions by all people seeking fair terms of cooperation. These we might think of as fundamental ‘relevant considerations’ or ‘reciprocal reasons’ in the context of NICE. A second set of relevant values concerns *which* non-effectiveness considerations are mutually justifiable – like a drug’s ability to provide life-extension at the end of life, a consideration for which NICE allows a higher cost-per-QALY threshold. These considerations are importantly linked to the contestation of QALYs: most claims about the QALY being an inappropriate measure – that, for instance, it subjects disabled people to ‘double jeopardy’³⁷⁹ – are claims about the types of other, non-effectiveness considerations that we should accept as relevant in distributive decisions.³⁸⁰ Although there is more disagreement in the literature about this set of values, I accept them for the purpose of this chapter as the ones that Appraisal Committees are obliged to consider as ‘relevant’ in making their decisions.

Rather than take issue with the relevancy of the principles set out by the NICE Board in *Social Value Judgements*, I consider here the judgements about applying these ‘relevant’ reasons that individual Technology Appraisal Committees must make in reaching a decision on a specific technology. By accepting at the outset of their deliberations that the values in *Social Value Judgements* are the set of relevant or mutually justifiable ones, it is the work of these Committees to reason in an accountable manner about how various arguments and types of evidence map onto these values and

³⁷⁹ John Harris, “QALYfying the Value of Life.,” *Journal of Medical Ethics* 13, no. 3 (1987); John Harris, “Double Jeopardy and the Veil of Ignorance--a Reply.,” *Journal of Medical Ethics* 21, no. 3 (1995).

³⁸⁰ Govind Persad, Alan Wertheimer, and Ezekiel J. Emanuel, “Principles for Allocation of Scarce Medical Interventions,” *Lancet* 373, no. 9661 (2009); Helen Mason, Michael Jones-Lee, and Cam Donaldson, “Modelling the Monetary Value of a QALY: A New Approach Based on UK Data,” *Health Economics* 18, no. 8 (2009).

bear upon their ordering or weighing in reaching a recommendation. The difficulty and persistence of disagreement on these judgements suggests, as I argued in Chapter III, that even if we agree on which reasons are ‘public,’ there is still a great deal of work to be done in specifying and interpreting these reasons when applying them within the context of a particular decision.

In this way, it is clear how the practical manifestation of the legitimacy problem in NICE is quite different from that in the HFEA. In the HFEA, the problem was a lack of clarity about how to reason in the arm’s length context and a corresponding failure to appropriately construct and justify procedural requirements. In NICE, the problem is that even when we have a reciprocal method of reasoning and attendant procedural requirements in place, when these have been at least provisionally settled in other words, questions remain about how to specify and interpret them in practice. In particular, we see continued confusion around the role of patients and the ‘patient voice’ in the appraisal process, particularly alongside the results of clinical trials and economic cost-effectiveness evaluations. Despite a commitment to accountability for reasonableness and agreement on the ‘relevant’ values that should form the normative basis of decisions, we do not always know how we should interpret and weigh different arguments and evidence in the Appraisal Committee context.

To resolve this tension, I argue in this chapter that the only way to involve patients in technology appraisals, given acceptance of accountability for reasonableness and the social values currently endorsed by NICE, is to treat their input as a type of evidence that can be relevantly compared to traditional forms of scientific and economic evidence. I outline the implications of this understanding of patient involvement as evidence provision for the *process* of involving patients in appraisals. When we cast patient input as evidence, however, we may continue to disagree about what this patient

evidence, as part of a larger body of evidence, entails for the *substantiation* of the guiding social value judgements. What does it mean to use evidence – particularly when its epistemic status is contested – in a publicly accountable way? To consider this I turn to a review of the NICE appeals process and its decisions to date. This process allows consultees – including patient groups, manufacturers, and primary care trusts – to contest the way in which the Appraisal Committee interpreted or weighed the evidence before it: one of the valid ‘grounds of appeal’ is the charge that a Committee’s decision was ‘unreasonable’ in light of the evidence submitted. In the terms of my framework, this ground of appeal allows appellants to argue that the way in which evidence was utilized was *non-reciprocal*, that its evaluation and use in forming a judgement was not based on standards or reasons that others could accept. I show through this review that although NICE has instituted a relatively strong appeals process, meeting the fourth condition of my framework – that of fostering the means of deliberative resistance – requires more from this process. This will help flesh out the meaning of this fourth condition as well as how it is distinct from Daniels’ and Sabin’s condition of challenge and revision.

1. Voice or choice: delimiting patient involvement in technology appraisals

Recall that in Chapter II, three key sources suggested the lack of a clear rationale for patient involvement in NICE technology appraisals. After hearing testimony from Laura Weir of Patients Involved in NICE that “there is no clear role for patient groups” in appraisals, the House of Commons Select Committee on Health advised, “It is important for the credibility of NICE and for the decisions that it makes that *the patient voice is effectively and openly represented* in all its work” (emphasis added).³⁸¹ The

³⁸¹ Health Select Committee, *National Institute for Clinical Excellence*, 16 January 2013, HC 782-I 2012-2013.

group *within* NICE responsible for supporting patient involvement, the Public Involvement Programme, has similarly emphasized the need for clarification about how patient experts “fit in” to appraisals.³⁸² And previous interviews by Timothy Milewa and Christine Barry, and by Pauline Quennell, conducted with both Appraisal Committee members and representatives of patient groups, illustrated a significant mismatch between these two groups of interviewees about the role that anecdotal and often emotive patient submissions or testimony plays, and should play, in appraisals.³⁸³ Given the goal of working through the fine tuning of a process based broadly on reciprocity, how can we make progress on this uncertainty about the place of patients in appraisals?

Traditionally, involving patients in the provision of healthcare has tended to evoke debate on whether such involvement is meant to grant patients ‘voice or choice,’³⁸⁴ whether they are to be ‘getting their *say*, or getting their *way*.’³⁸⁵ On closer inspection, however, this dichotomy between ‘patient voice’ and ‘patient choice’ is not particularly fruitful as a way of thinking about the role of patient involvement within NICE technology appraisals. To see why this is the case, consider Sherry Arnstein’s ‘ladder of citizen participation,’ which has been highly influential in attempts to theorize both patient and public involvement. Arnstein classifies different types of citizen

³⁸² Amis and Livingstone, “Technology Appraisal Patient Expert Survey 2012 Report.”

³⁸³ Milewa and Barry, “Health Policy and the Politics of Evidence”; Milewa, “Health Technology Adoption and the Politics of Governance in the UK”; Pauline Quennell, “Getting Their Say, or Getting Their Way? - Has Participation Strengthened the Patient ‘Voice’ in the National Institute for Clinical Excellence?”; Pauline Quennell, “Getting a Word in Edgeways? Patient Group Participation in the Appraisal Process of the National Institute for Clinical Excellence.”

³⁸⁴ For accounts of patient (and public) involvement that center on the voice/choice dichotomy, see Rob Baggott, “A Funny Thing Happened on the Way to the Forum? Reforming Patient and Public Involvement in the NHS in England,” *Public Administration* 83, no. 3 (2005); Rudolf Forster and Jonathan Gabe, “Voice or Choice? Patient and Public Involvement in the National Health Service in England under New Labour,” *International Journal of Health Services* 38, no. 2 (2008); David Hughes, Caroline Mullen, and Peter Vincent-Jones, “Choice vs. Voice? PPI Policies and the Re-Positioning of the State in England and Wales,” *Health Expectations: An International Journal of Public Participation in Health Care and Health Policy* 12, no. 3 (2009).

³⁸⁵ Pauline Quennell, “Getting Their Say, or Getting Their Way? - Has Participation Strengthened the Patient ‘Voice’ in the National Institute for Clinical Excellence?”

participation based on the degree to which they grant ‘citizen power.’³⁸⁶ Rungs on the ladder representing a dearth of participation (such as public education or information programmes) or rungs representing mere ‘window dressing participation’ (such as surveys, consultations, and token appointments to public bodies), are lower on the ladder precisely because they fail to give citizens any real decision-making power. Giving patients a ‘voice’ in policymaking on this account is distinct from giving them real ‘choice,’ Arnstein says, because “under these conditions they lack the power to ensure that their views will be heeded by the powerful...there is no follow through, no ‘muscle,’ hence no assurance of changing the status quo.”³⁸⁷

Patient participation as patient choice, then, requires the muscle or power to affect the decisions that patients want. It can perhaps *only* be distinguished from mere window dressing participation by being able to explicitly point to the instances and ways in which patients get what they want. Technology Appraisal Committees exist, however, because limited healthcare resources mean that patients cannot always get the treatments they want, or indeed need. There will be times when even clinically effective new drugs are too expensive to fund, given the opportunity cost to the rest of the NHS and other citizens. It is because of this collective dimension that ‘choice’ in the context of technology appraisals should be driven by considerations that all citizens – not merely specific patient groups, or indeed any other interested parties like pharmaceutical companies or doctors – can accept. Procedurally, this is secured by ceding the identification of those considerations to the Citizens Council and the NICE Board, which are then made explicit in *Social Value Judgements*. Indeed, the fifth principle of *Social Value Judgements* states that demands for patient choice must be tempered by

³⁸⁶ Sherry R. Arnstein, “A Ladder Of Citizen Participation,” *Journal of the American Institute of Planners* 35, no. 4 (1969).

³⁸⁷ *Ibid.*, 217.

considerations of clinical effectiveness, cost effectiveness, and other (mutually justifiable) social values.

Speaking of ‘voice or choice’ in the context of technology appraisals thus seems to be a red herring, as the only plausible account of patient involvement in appraisal decisions is that of having one’s ‘voice’ heard. But, even if we can move past the voice or choice dichotomy, what does it mean to have one’s voice *heard*? Or, framed with procedural requirements in mind, what processes must be put in place to ensure that the patient’s voice is heard? On the one hand, this could mean simply giving patients a space in which to voice their perspective and attempt to persuade the Committee. This understanding of accounting for the patient voice suggests that we meet our duties to patients or other special interest groups affected by a decision when we give them the opportunity to register their comments or have a ‘patient expert’ represent their views before an Appraisal Committee.

If the ‘patient choice’ interpretation of involvement is too strong within a public accountability framework guided by reciprocity, this interpretation of the patient voice seems far too weak. It simply requires that we provide the opportunity for patients to speak while providing them with no assurance that they will be heard. Many have noted how, “in the absence of real influence, the illusion of voice can lead to even greater frustration and disenchantment than having no voice at all.”³⁸⁸ The challenge, then, is to situate patient involvement within the deliberative process such that the ‘real influence’ of the patient voice is demonstrable even when patient groups advocating for a positive recommendation of a technology do not get the outcome they want. We must be able to show that patient views are “heeded,” in Arnstein’s terms, even when these views are

³⁸⁸ Michael X. Delli Carpini, Fay Lomax Cook, and Lawrence R. Jacobs, “Public Deliberation, Discursive Participation, and Citizen Engagement: A Review of the Empirical Literature,” *Annual Review of Political Science* 7, no. 1 (2004).

ultimately defeated by other considerations, like cost. Reciprocity suggests that to do this, we need to be able to give reasons that patients can recognize as relevant for *why* other considerations prevail when they do. To the extent that patient groups fail to recognize the public values set out in *Social Value Judgements*, by merely claiming – for example – that rationing is morally wrong or that their clinical need necessitates a drug’s recommendation, responding to the patient voice will simply require clear, public reference to those mutually justifiable reasons.

But often, patient groups accept and argue on the basis of these public reasons, making claims about the extent to which a drug is clinically effective or what the true costs of a treatment, or its main comparator, are. In such cases, where patients present arguments based upon values that are clearly *relevant*, in the way demanded by reciprocity, we must be able to adjudicate between their claims and the claims made by scientists or economists about these same values. To the extent that we continue to conceive of patient involvement as disconnected from and based upon a different rationale than the ‘technical’ data related to clinical and cost effectiveness, viable comparison between the patient voice and clinical data or economic models will not be possible. To be commensurable, both must be understood and treated as forms of *evidence* that have the opportunity to actually impact the Appraisal Committee’s interpretation or weighing of social values. This properly *constrains* the role of patients and patient groups to providing evidence that affects the interpretation or specification of these values, as the values themselves should be based on reasons that all fair-minded citizens – not merely specific patient groups – can accept, and thus should be set with reference to public involvement through the Citizens Council. Yet this move also *strengthens* the role of patients by giving them a genuine and justifiable opportunity to impact decisions.

What does it mean to take seriously patient input as evidence? Currently, as set out in Chapter II, the formal avenues for patient group involvement in technology appraisals include consulting on the scope of the appraisal, submitting ‘written evidence’ for the Committee to consider, and nominating a patient expert to attend the Committee meeting to provide ‘oral evidence.’ If we accept that the role of patients in technology appraisals should be as evidence providers, do we need to reassess or adjust any of these processes for involving them?

2. The patient voice as evidence

Without diving too deep into the extensive debate about what ‘evidence’ is,³⁸⁹ it may be broadly defined as information that is helpful in forming a conclusion or judgement by indicating whether a proposition is true or valid.³⁹⁰ In the context of an Appraisal Committee, evidence can thus be understood as information that is helpful in determining whether or not to recommend a treatment: in determining what attention to the public reasons set out in *Social Value Judgements* requires within the context of a *particular* case. To think about how patients can provide such information, we first need a more general understanding of the other types of evidence that NICE considers in making these judgements and how Appraisal Committees evaluate the strength and weight of these different types of evidence.

In appraising a new treatment for a given condition and patient population, a Committee must compare it to alternative treatments, called comparators, with the goal of determining whether a difference in expected costs between the treatment and its

³⁸⁹ This may be better characterized as multiple debates, as the nature of evidence is contested somewhat insularly across a range of fields, including philosophy, law, medicine, and science.

³⁹⁰ “Evidence,” *Oxford Dictionaries* (Oxford University Press), last accessed August 14, 2015, <http://www.oxforddictionaries.com/definition/english/evidence>.

comparators is justified by changes in expected health effects. Costs are those to the NHS for the drug or treatment itself as well as other resources associated with its use (like hospital stay), while health effects include both changes in the length and quality of life, expressed as QALYs.

The measurement of life-year health effects is fairly uncontroversial: all else equal, if patients in a trial survive for an average of four years on Drug A and six years on Drug B, we can calculate that the use of Drug B will produce an expected two additional life-years. Assessing health-related quality of life is much more difficult, because we need to determine 1) what factors count in health-related quality of life, 2) how those factors are affected by a treatment and its comparator, and 3) how we should *value* changes in those factors. For instance, to assess a drug that reduces pain in a chronic condition, we would first need to identify pain as a relevant factor in quality of life, then measure the difference in pain levels between patients on a treatment and ones on a comparator, and finally – assuming the drug reduces pain better than the comparator – quantitatively determine how much we value that reduction in pain. NICE notes, “Health-related quality of life, or changes in health-related quality of life, should be measured directly by *patients*,” but the value of these changes must be based on *public* preferences from a representative sample of the UK population (emphasis added).³⁹¹

To standardize measurements of quality of life, NICE’s preferred measure is the EQ-5D, a patient questionnaire used frequently in trials to document patients’ quality of life.³⁹² It identifies five factors that count in health-related quality of life: 1) mobility, 2) ability to self-care, 3) ability to undertake usual activities, 4) pain and discomfort, and 5) anxiety and depression. Patients may choose one of three ‘levels of severity’ on each

³⁹¹ National Institute for Health and Care Excellence, Section 5.3 Measuring and Valuing Health Effects, “Guide to the Methods of Technology Appraisal 2013.”

³⁹² Ibid.

factor: 1-no problems, 2-some problems, or 3-severe problems. So results from a patient who reports experiencing ‘some problems’ in each of the five factors would be coded as ‘health state’ 22222. Separately, in a representative sample of the public, participants are faced with a hypothetical situation in which they have 10 years to live at health state 22222. They are then asked how much time they would be willing to give up to live at full health (health state 11111, no problems across all five factors in health-related quality of life). Averaging and extrapolating these ‘time trade-off’ responses provides the *value* of moving from one health state to another. This method means that patients themselves do not determine the value of changes in quality of life. Although it is recognized, for instance, that an athlete may value health gains in mobility more than the average person, the idea is that to treat all NHS citizens fairly, NICE must be able to compare the value of health gains across treatments and people. Nevertheless, as I will illustrate shortly, patients *are* crucially important in the first two steps of identifying the full range of factors that affect health-related quality of life and reporting how different treatments affect these factors.

In assessing effects on both length and quality of life, “The Institute has a preference for RCTs directly comparing the intervention with 1 or more relevant comparators.”³⁹³ Randomised controlled trials are traditionally considered the ‘gold standard’ in measuring clinical benefit in evidence-based medicine, with systematic reviews of RCTs and meta-analyses of RCTs often represented at the top of the ‘evidence hierarchy,’ followed by non-randomised and non-controlled studies (like observational studies), and finally case reports and expert opinion at the bottom.³⁹⁴

³⁹³ National Institute for Health and Care Excellence, Section 5.2 Synthesis of Evidence on Health Effects, “Guide to the Methods of Technology Appraisal 2013.”

³⁹⁴ OCEBM Levels of Evidence Working Group, “OCEBM Levels of Evidence” (Oxford Centre for Evidence-Based Medicine), last accessed July 17, 2015, <http://www.cebm.net/index.aspx?o=5653>; Jeremy

Evidence from RCTs is typically afforded this weight because this method is capable of providing a high degree of internal validity: because participants are randomised and there is a control group with which to compare the treatment group, we can be more sure that any effects we see in the treatment group are the result of the treatment itself as opposed to selection bias or other confounding factors.

However, the limits of RCTs are now well recognized, both for evidence-based medicine as a whole³⁹⁵ and within the context of health technology appraisal specifically.³⁹⁶ Two limits are particularly problematic for cost effectiveness analyses. First, there is a growing recognition that the quality of life factors identified by researchers as end points within clinical trials are not always reflective of the factors – or at least the totality of factors – that patients themselves recognize as influential in health-related quality of life. The attention granted to this issue owes much to the patient-centred care movement, which champions the observation that what patients actually care about is not always obvious and may differ significantly from what clinicians or researchers *think* patients care about (or think patients *should* care about).³⁹⁷

Of course, many factors that affect patients' quality of life will be ones that doctors and researchers have always attended to. For instance, the incidence of increased suicidal thoughts and tendencies as a side effect of an anti-depressant is an outcome that obviously matters to patients, but it is also one that even the most traditional doctor-centred models of medicine would recognize as a key outcome due to its relationship to

Howick et al., "The 2011 Oxford CEBM Levels of Evidence (Introductory Document)" (Oxford Centre for Evidence-Based Medicine), <http://www.cebm.net/index.aspx?o=5653>.

³⁹⁵ Alvan R Feinstein and Ralph I Horwitz, "Problems in the 'Evidence' of 'Evidence-Based Medicine,'" *The American Journal of Medicine* 103, no. 6 (1997).

³⁹⁶ Nick Black, "Why We Need Observational Studies to Evaluate the Effectiveness of Health Care," *BMJ* 312, no. 7040 (1996); David Evans, "Hierarchy of Evidence: A Framework for Ranking Evidence Evaluating Healthcare Interventions," *Journal of Clinical Nursing* 12, no. 1 (2003); Mark Petticrew and Helen Roberts, "Evidence, Hierarchies, and Typologies: Horses for Courses," *Journal of Epidemiology and Community Health* 57, no. 7 (2003).

³⁹⁷ Moira Stewart, "Towards a Global Definition of Patient Centred Care," *BMJ* 322, no. 7284 (2001).

mortality rates. However, the relative importance of another much discussed outcome, sexual side effects, is not intuitively obvious nor does good evidence exist on the question. As a group of health researchers noted in calling for attention to this question, “Whilst most of the research [on the sexual side effects of selective serotonin reuptake inhibitor (SSRI) medication] has focused on the prevalence and treatment of sexual difficulties, little is known about how patients cope with the SSRI-related sexual side effects.”³⁹⁸ Is loss of libido the least of concerns for patients on anti-depressants or is it a central concern greatly affecting their quality of life that needs to be included as a quality of life factor to be measured in trials?

A prime example of the gap between what patients care about and what researchers think patients care about can be found in NICE’s technology appraisal of etanercept and infliximab, two treatments for psoriasis (TA104).³⁹⁹ The appraisal initially relied heavily upon the Psoriasis Area Severity Index (PASI) in its clinical and cost effectiveness measurements, as PASI “is the most widely used measurement tool for psoriasis in clinical trials.”⁴⁰⁰ PASI scores are based on the severity and total area affected by psoriasis flare-ups, and were thought to be the most important factor in measuring quality of life improvements. However, during the consultation process, patients brought it to the attention of the Appraisal Committee that it is the *location* on the body rather than the *size* of the outbreak that matters most.⁴⁰¹ Patients indicated that quality of life is improved more significantly by a drug that reduces outbreaks on their face or other very exposed areas than one that provides a greater reduction in total

³⁹⁸ Cathy O’Mullan et al., “Women’s Experiences of Coping with the Sexual Side Effects of Antidepressant Medication,” *Psychology & Health* 29, no. 12 (2014).

³⁹⁹ National Institute for Health and Care Excellence, “Etanercept and Infliximab for the Treatment of Psoriatic Arthritis,” *Guidance*, July 2006, <https://www.nice.org.uk/guidance/ta104>.

⁴⁰⁰ National Institute for Health and Care Excellence, “Etanercept and Efalizumab for the Treatment of Adults with Psoriasis: Post-Appeal Final Appraisal Determination (TA103),” June 9, 2006, <http://www.nice.org.uk/guidance/TA103/documents/postappeal-final-appraisal-determination-psoriasis-efalizumab-and-etanercept>, 3.

⁴⁰¹ Amis, “Patient Involvement in NICE Technology Appraisals,” 37.

affected area but that targets the back or less visible locations. This suggests a kind of ‘experiential knowledge’ that patients possess that constitutes a valid type of evidence.⁴⁰²

On the grounds that doctors and researchers do not have access to this knowledge, as evidenced by the psoriasis example and despite their other important expertises, the ability to provide this knowledge can be interpreted as a form of ‘patient expertise.’⁴⁰³

The first limitation of evidence from RCTs, that the end points considered do not necessarily reflect the factors that affect patients’ quality of life, is thus a crucial point at which evidence from patients and patient groups can help inform an Appraisal Committee’s attempt to interpret and weigh the reciprocal values of clinical effectiveness and cost. The second key limitation of evidence from RCTs for technology appraisal is that, to the extent that patients who enrol in clinical trials are not representative of the total patient population, the evidence on health effects from even otherwise well-designed trials (ones that, for instance, used patient-oriented end points) will be biased. On this problem of external validity, philosopher Nancy Cartwright explains, “A properly conducted RCT provides evidence that the intervention works somewhere (i.e. in the trial). The decision maker, however, needs to estimate ‘will it work for us?’”⁴⁰⁴ Differences in both the social, psychological, and physical facts of the trial, on the one hand, and the target population and real clinical practice, on the other, mean that we must be careful in moving from ‘it works somewhere’ claims to ‘it will work for us’ claims.⁴⁰⁵

⁴⁰² Jo Rycroft-Malone et al., “What Counts as Evidence in Evidence-Based Practice?,” *Journal of Advanced Nursing* 47, no. 1 (2004).

⁴⁰³ Sally E Thorne, Kerstin Ternulf Nyhlin, and Barbara L Paterson, “Attitudes toward Patient Expertise in Chronic Illness,” *International Journal of Nursing Studies* 37, no. 4 (2000); Ian Kennedy, “Patients Are Experts in Their Own Field,” *BMJ* 326, no. 7402 (2003); Vikki A. Entwistle et al., “Lay Perspectives: Advantages for Health Research,” *BMJ* 316, no. 7129 (1998).

⁴⁰⁴ Nancy Cartwright and Eileen Munro, “The Limitations of Randomized Controlled Trials in Predicting Effectiveness,” *Journal of Evaluation in Clinical Practice* 16, no. 2 (2010): 265.

⁴⁰⁵ Nancy Cartwright, “A Philosopher’s View of the Long Road from RCTs to Effectiveness,” *The Lancet* 377, no. 9775 (2011).

To illustrate this limitation of RCTs, consider an example discovered during my review of NICE's appeals decisions (discussed at greater length in section 3). In 2007, NICE appraised and initially chose not to recommend erlotinib for non-small cell lung cancer (TA162), on the basis that it was not a cost-effective use of resources relative to its comparator docetaxel, a type of chemotherapy treatment.⁴⁰⁶ In coming to this decision, the Appraisal Committee relied upon data from an RCT that the rate of febrile neutropenia on docetaxel was 1.8%.⁴⁰⁷ Febrile neutropenia is the development of a fever in the presence of a reduced white blood cell count caused by chemotherapy treatment, a serious and costly effect that can lead to death because patients have trouble fighting off infection with this reduced count. Because erlotinib is a kinase inhibitor rather than a chemotherapy drug, it is not associated with this effect.

On appeal, two patient groups, Roy Castle Lung Cancer Foundation and Cancerbackup, brought evidence from the most recent American Society of Clinical Oncology meeting that the average rate of febrile neutropenia across trials presented was 6.6%, much higher than the incidence of 1.8% accepted by the Appraisal Committee.⁴⁰⁸ The true rate of febrile neutropenia, they suggested, is not reflected by that seen in the RCT. Additionally, the manufacturer of erlotinib and relevant professional societies used both the testimony of a clinical expert and reference to published clinical opinion to suggest that the true *cost* of this side effect was significantly higher than estimated in the cost effectiveness models. Together, these two pieces of evidence convinced the Appeal Panel that the Appraisal Committee should reconsider the guidance using a "wider range

⁴⁰⁶ National Institute for Health and Care Excellence, "Erlotinib for the Treatment of Non-Small-Cell Lung Cancer," *Guidance*, November 2008, <https://www.nice.org.uk/guidance/ta162>.

⁴⁰⁷ National Institute for Health and Care Excellence, "Decision of the Appeal Panel: Guidance on Erlotinib for the Treatment of Non-Small Cell Lung Cancer," June 6, 2007, <https://www.nice.org.uk/guidance/ta162/documents/appeal-decision2>, 18.

⁴⁰⁸ *Ibid.*

of realistic values for the incidence and costs of febrile neutropenia,” given that these could alter the recommendation.⁴⁰⁹

This example shows how the difference between trial rates and clinical practice rates can truly mean the difference between a drug being cost effective and cost ineffective. To the extent that patient groups can help suggest relevant differences between participants in RCTs and the UK patient population, based on their connection to and ability to mobilize this broader population, they can provide information that usefully addresses this limitation of RCTs.

However, if these two limitations of RCTs suggest ways in which patient input can function as evidence, they also suggest key challenges for the patient voice. I will show how this recognition means that some processes of patient involvement will be better than others: that some methods provide more appropriate information for a Committee attempting to reason reciprocally. In particular, the function of the patient voice in technology appraisals faces two key hurdles: one of representativeness and a second of integrating quantitative and qualitative evidence. With respect to the first challenge, patient experts providing oral testimony before an Appraisal Committee “are asked to consider whether their condition and the problems and benefits of the technology in question are accurately represented,” as well as “whether the quality-of-life measures in clinical trials are reflective of what matters to patients.”⁴¹⁰ Although evidence on these questions is precisely what we need to remedy the limitations of trial data, it is far from clear that patient experts can provide it. Patients are ‘expert,’ as suggested above, with respect to their *own* experiential knowledge. While the *systematic* aggregation and analysis of patient views on quality of life factors may provide an

⁴⁰⁹ Ibid., 28.

⁴¹⁰ NICE Public Involvement Programme, “Hints and Tips for Nominating Patient Experts for a Single Technology Appraisal (STA).”

‘accurate representation’ of patient perspectives, patient experts have no claim to representativeness in this respect, even if they see themselves as representatives of the ‘patient interest.’ Even a carer serving as a patient expert who has seen a representative sample of patients will struggle to provide any support for this representativeness beyond her own word.

Although we want the patient voice to bear upon technology appraisals, we do not want it to be at the expense of the central principles of evidence-based medicine. Individual patients, while perhaps useful as *members* of an Appraisal Committee in challenging disciplinary ways of thinking (see Chapter IV on lay representation), will continue to be subject to accusations that the information they provide as ‘evidence’ is biased, subjective, or anecdotal, as levelled by the members of Appraisal Committees interviewed by Milewa and Barry.⁴¹¹ Indeed, the patient expert as the main method of patient involvement risks institutionalizing the ‘weak’ account of the patient voice discussed in the previous section, in which patients may not be ‘heard’ due to the incompatibility of their voice with the methods of appraisal: as health economists John Bridges and Christopher Jones note, “Although many [health technology appraisal] agencies have reacted to [the] call to become more patient centered, this action has often been limited to giving ‘professionalized’ patients a seat, but quite often not an adequate voice, at the decision-making table.”⁴¹²

Written submissions from patient interest groups are meant to represent the larger patient population. Indeed, to the extent that such groups can leverage their institutional position and buy-in from patient supporters to coordinate systematic collection of patients’ perspectives, this represents a better method of patient involvement than

⁴¹¹ Milewa and Barry, “Health Policy and the Politics of Evidence.”

⁴¹² John F. P. Bridges and Christopher Jones, “Patient-Based Health Technology Assessment: A Vision of the Future,” *International Journal of Technology Assessment in Health Care* 23, no. 1 (2007).

individual patient experts, one that is more clearly tailored to the purpose of evidence provision. But, as we will see in the following section, they similarly struggle to demonstrate the reliability of their evidence. Various groups have attempted to collect data through quantitative and qualitative surveys of their patient-members, but the design is frequently underdeveloped and the assumption of bias is strong when carried out by the interest group itself. The Health Technology Assessment international (HTAi) interest group on patient/citizen involvement has argued that health agencies like NICE should commission independent primary research to access this information from patients about quality of life factors or other experiences if there is no good existing evidence on it.⁴¹³ While this is methodologically ideal, it will often be too expensive and time-consuming; its importance must be weighed against the need to make decisions about life-saving treatments quickly and against broader cost concerns.

The second challenge that patient-sourced evidence faces is that of comparing qualitative and quantitative data. For instance, an Appraisal Committee may be convinced that a robust and independent qualitative study of patient experiences within the UK provides good evidence that certain quality of life factors were not captured in the RCT or other clinical data. This removes the representativeness problem that patient expert testimony and patient group submissions face. However, the Committee still has to decide how to incorporate this finding into the quantitative clinical and cost effectiveness models that it uses to make and justify its final decision. Although some general methodological progress has been made on synthesizing qualitative and

⁴¹³ Karen Facey et al., “Patients’ Perspectives in Health Technology Assessment: A Route to Robust Evidence and Fair Deliberation,” *International Journal of Technology Assessment in Health Care* 26, no. 3 (2010).

quantitative health data, these methods are not particularly conducive to the output of a single measure, like cost per QALY or incremental cost effectiveness ratios.⁴¹⁴

More generally, in bringing together quantitative and qualitative evidence, NICE notes:

[I]t is acknowledged that the concept of validity in qualitative research is less clearly defined than for quantitative research. As a result, the review team may wish to take account of other factors when judging the ‘trustworthiness’ of the [qualitative] study, its relevance to the research questions and how ‘convincing’ the results are.⁴¹⁵

Making these judgements about the convincingness of evidence and how it weights up to other *types* of evidence from different *sources* is difficult. Annabelle Lever attributes this to our current lack of a ‘political theory of evidence.’⁴¹⁶ Given that we “have to decide the relative quality and weight of evidence from different sources, created by different methodologies, the relative merits of which may be hard to determine”:

...further reflection is desirable on what is to count as evidence in the justification of collectively-binding decisions, and of how that evidence is to be used and weighed by decision makers, who are rarely in a position to appraise its epistemological merits for themselves, and will often receive reports of what others have discovered only at second or third hand.⁴¹⁷

We often disagree or are simply uncertain about the standards for evidence, its comparison, or its interpretation in the technology appraisal setting. This is troublesome for reciprocal reasoning, which requires not only the *values* we base collective, coercive

⁴¹⁴ Nicholas Mays, Catherine Pope, and Jennie Popay, “Systematically Reviewing Qualitative and Quantitative Evidence to Inform Management and Policy-Making in the Health Field,” *Journal of Health Services Research & Policy* 10, no. 1 (2005); Mary Dixon-Woods et al., “Synthesising Qualitative and Quantitative Evidence: A Review of Possible Methods,” *Journal of Health Services Research & Policy* 10, no. 1 (2005).

⁴¹⁵ National Institute for Health and Care Excellence, Section 5.3 Assessing the Quality of the Evidence, “Methods for the Development of NICE Public Health Guidance (third Edition),” September 2012, <https://www.nice.org.uk/article/pmg4/chapter/reviewing-the-scientific-evidence>.

⁴¹⁶ Cartwright and Munro, “The Limitations of Randomized Controlled Trials in Predicting Effectiveness.”

⁴¹⁷ Annabelle Lever, “Democracy and Lay Participation: The Case of NICE,” in *The Future of Public Service Reform*, ed. Henry Kippin Gerry Stoker (London: Bloomsbury Academic Press, 2013).

decisions upon but also the standards of evidence to be ones that reasonable people can accept. And, even if we can agree on the general standards, reasonable people may still disagree about what the available evidence within discrete cases entails for these standards, as suggested by the ‘burdens of judgement’ discussed in Chapter III. Turning to evidence to solve health governance disagreements will not calculate an unproblematic answer, in other words, if we disagree about what the evidence base *means*.⁴¹⁸

This creates a practical challenge to reciprocal reasoning in that, under such circumstances, we may struggle to explain why the evidence weighs in favour of one outcome over the other, and how we find certain sources of evidence more robust or more convincing than others. I have argued that substantiating NICE’s social value judgements (understood as a set of reciprocal reasons) in practice requires treating patients as providers of evidence within the appraisal process, as well as supporting patient involvement methods that are fit for this purpose. Yet this useful development on the confusion surrounding the role of the patient voice – the initial practical challenge set out in Chapter II – exacerbates this second practical challenge of public reasoning about *evidence*. Committees must compare qualitative patient evidence, which may be more or less representative, to quantitative but recognizably flawed RCTs and economic models in the absence of agreement (even amongst experts) about how to manage such a synthesis.

NICE’s current method for adjudicating disputes about what the ‘reasonable’ interpretation of evidence demands is its appeals process. Within this process, which is

⁴¹⁸ On the politics and contested nature of ‘evidence’ in evidence-based health policy, see R.E.G. Upshur, E.G. Vandenkerkhof, and V. Goel, “Meaning and Measurement: An Inclusive Model of Evidence in Health Care,” *Journal of Evaluation in Clinical Practice* 7, no. 2 (2001); Maya J. Goldenberg, “On Evidence and Evidence-Based Medicine: Lessons from the Philosophy of Science,” *Special Issue: Gift Horse or Trojan Horse? Social Science Perspectives on Evidence-Based Health Care* 62, no. 11 (2006); Nick Black, “Evidence Based Policy: Proceed with Care,” *BMJ* 323, no. 7307 (2001).

the embodiment of Daniels' and Sabin's third condition of accountability for reasonableness, patient groups and other consultees can appeal the decision of an Appraisal Committee. One of three valid grounds of appeal is that the Committee prepared guidance that is 'unreasonable' in light of the evidence it considered. Although each technology appraisal officially stands alone and future Committee decisions need not reflect the same considerations as past ones, these decisions create unofficial precedent within NICE that is often cited in future decisions. As such, the approach to evidence in NICE technology appraisals and the standards for its reasonable interpretation are heavily impacted by this appeals process. In the next section, I draw on the results of a review that I undertook of all 86 technology appraisal decisions that were appealed between 1 March 2000 (the first NICE appraisal) and 31 March 2015. While I will present some more general results of the review for context, its role in this chapter is to allow for the analysis and critique of NICE's processes for both initially collecting and then weighing evidence. Although I find the appeals process to be a relatively robust interpretation of Daniels' and Sabin's condition of challenge and revision, I suggest certain shortcomings of its responsiveness to contested evidence on the basis of my framework's fourth condition of fostering the means of deliberative resistance. This usefully allows me to flesh out this distinction between our approaches to reciprocity, while also making progress on what reciprocal reasoning about contested evidence requires.

3. Contesting evidence: fostering the means of resistance in NICE appeals

Between 1 March 2000 and 31 March 2015, NICE published a total of 338 technology appraisals resulting in 578 individual recommendations (as some appraisals

consider multiple technologies at once).⁴¹⁹ As such, the 86 appeals that I reviewed represent only approximately 25 percent of the total technology appraisals published by NICE. Why focus on this subset of decisions? First, as suggested in the last section, the appeals process is a central site at which the ‘reasonable’ use of evidence within NICE is contested, evaluated, and revised, and it creates a body of precedent to which Appraisal Committees refer. A second reason to focus on appraisals that have been appealed is the nature of cases that make it to this point of the appraisal process. As only a quarter of appraisal determinations are appealed, examining them will not give a representative picture of the average appraisal process. NICE has positively recommended technologies, either within the broader patient population or within specific subgroups, in 80 percent of its 553 recommendations.⁴²⁰ Except in a few cases in which clinical commissioning groups challenge the recommendation of a technology based on cost considerations, the overwhelming majority of appealed cases are drawn from the 20 percent in which NICE did *not* recommend the technology in question. As such, unlike ‘easier’ cases where the clinical and cost effectiveness considerations align with the patient input in supporting the recommendation, cases that are appealed tend to be those in which the two types of evidence appear to conflict.

Consultees can appeal on one of three⁴²¹ grounds:

Ground 1: The Institute has failed to act fairly in accordance with the published procedures as set out in the Institute’s Guide to the Technology Appraisal Process;

Ground 2: The Institute has prepared guidance that is perverse in light of the evidence submitted;

Ground 3: The Institute has exceeded its legal powers.

⁴¹⁹ National Institute for Health and Care Excellence, “NICE Statistics,” <https://www.nice.org.uk/news/nice-statistics>.

⁴²⁰ Ibid.

⁴²¹ In 2013, these three grounds were folded into two: “Ground 1: In making the assessment that preceded the recommendation, NICE has: 1a. failed to act fairly, or 1b. exceeded its power. Ground 2: The recommendation is unreasonable in light of the evidence submitted to NICE.” In my review, however, I reclassified appeals after 2013 using the three-ground framework to keep results standardized.

After 2013, Ground 2 was updated to read, “The recommendation is *unreasonable* in light of the evidence submitted to NICE,” rather than ‘perverse.’ From the perspective of accountability for reasonableness and the challenge/revisions condition, Grounds 1 and 2 thus allow for two distinct challenges about NICE’s reasonableness: Ground 1 focuses on *procedural* unreasonableness, that a Committee failed to adhere to the fair process NICE has created and by which it is bound, while Ground 2 deals with *substantive* unreasonableness, that the substance of the decision itself is not reasonable. NICE thus recognizes that a decision can be unreasonable even if it is the result of a fair process.

This returns to my claim in Chapter III that Daniels’ and Sabin’s ‘procedural’ condition of relevancy in fact is based on a substantive account of public reason, given that it states, “[A] rationale will be reasonable if it appeals to evidence, reasons, and principles that are accepted as relevant by fair-minded people who are disposed to finding mutually justifiable terms of cooperation.”⁴²² When Appraisal Committees must make decisions where fair-minded people cannot agree on the relevancy of evidence, standards for evidence, or what uncertainty in evidence implies – as the burdens of judgement suggest is likely – it may be that there exist multiple ‘reasonable’ judgments. This was described in Chapter III as the problem of inconclusiveness in public reason. In such instances, although it is not the case that the Committee makes a decision that is *unreasonable* in light of the evidence, it may be the case that other reasonable decisions existed. I argued in that chapter that reciprocity demands that a body like NICE must do more than show that its decision is not unreasonable in such cases: it must take an extra step to foster the means of deliberative resistance for those who have reasonable grounds

⁴²² Daniels and Sabin, *Setting Limits Fairly: Can We Learn to Share Medical Resources?*, 45.

for disputing the conclusion. I hope to show in this analysis of the appeals process exactly what that extra step entails.

A few more descriptive elements of the appeals process are necessary first, however. In April 2004, NICE introduced an ‘initial scrutiny’ stage of the appeal process, to “confirm that the points of appeal put forward are at least arguably within the permitted grounds of appeal.”⁴²³ Prior to that, there was no formal mechanism for reviewing the grounds of appeal. In 2006, NICE began publishing the letters of appeal received from consultees, and by 2007 it was also consistently publishing its scrutiny letters in response to those letters of appeal. This means that the appeals process now essentially has two stages, both of which may be publicly inspected. The scrutiny letters shine light on the institution’s gatekeeping of the appeal process, as only those points of appeal that are determined to fall within one of the valid grounds of appeal may pass. And the written decision of the Appeal Panel demonstrates how it reasons about whether to uphold or dismiss those points of appeal.

Assuming that there is at least one valid point of appeal, an Appeal Panel is formed. The Panel is composed of five members: the chair, who is an external member and either a healthcare provider or a patient representative, one non-executive director of NICE, one NHS representative, one representative from industry, and one patient representative. Appellants may send up to five representatives before the Panel, including legal representation. After brief opening statements by appellants, the Panel may ask any questions it deems relevant. While the exact minutes of the hearing are not published (although members of the public may register to attend), the written decision of the Panel includes a detailed account of the questions and responses at the hearing.

⁴²³ National Institute for Health and Care Excellence, Section 5: Scrutiny of Appeals Points, “Guide to the Technology Appraisal and Highly Specialised Technologies Appeal Process,” February 2014, <https://www.nice.org.uk/article/PMG18/chapter/Foreword>. The first appeal to use the new initial scrutiny stage was heard in February 2005 (TA90).

The Panel's judgment may take one of three forms. First, the Panel may uphold one or more points of appeal and return the appraisal to the Appraisal Committee with instructions to reconsider certain components. Second, the Panel may uphold one or more points of appeal but find that it is unnecessary for the Appraisal Committee to reconsider it. This is usually if minor changes in wording are required. Finally, the Panel may dismiss all points of appeal, in which case the Final Appraisal Determination is published as guidance. There is no further option for appeal within NICE, although in two cases pharmaceutical companies have taken NICE to the High Court on judicial review (*R (on the application of Eisai Ltd) v NICE* in 2008 on Alzheimer's drugs and *Servier Laboratories v NICE* in 2010 on osteoporosis treatment).

To review these appeal decisions, I began with NICE's published list of all past appeals. I used publicly available documents – letters of appeal, letters of scrutiny of appeal points, and the decisions of the Appeal Panel, to determine in each case:

1. Which types of consultee filed an appeal (pharmaceutical company, patient/carer interest group, health professional society, or primary care trust/clinical commissioning group).
2. Upon which grounds of appeal each consultee filed.
3. Which grounds of appeal were heard by the Appeal Panel.
4. Which grounds of appeal were upheld and which were dismissed in each case.

After extracting this basic information, I focused on cases in which Ground 2 points of appeal were brought, particularly those brought by patient interest groups. In such cases, I attempted to pull out the central claims made about the unreasonableness of the consideration of the evidence and how the Appeal Panel addressed those claims. This allowed me to identify the main ways in which patient groups and the evidence that they provide or contest have functioned, when and why they have had their appeals about the unreasonableness of the interpretation of evidence upheld, and how – if their appeals are dismissed – NICE justifies that dismissal.

Categories of Appellant	Number of appeals in which category of appellant was represented
Pharmaceutical companies	70
Patient/carer interest groups	39
Health professional societies	33
PCTs/CCGs	6

Table 3. Types of appellants represented in March 2000 – March 2015 appeals.

Ground of Appeal ⁴²⁴	Rejected on scrutiny	Dismissed on appeal	Upheld on appeal
1 – Procedurally unfair	7	47	22
2 – Unreasonable in light of evidence	4	53	22
3 – Exceeded power	10	29	1

Table 4. Grounds of appeal invoked in March 2000 – March 2015 appeals.

In addition to a large group of Ground 1 appeals having to do with failure to follow process (such as failing to consult on the scoping stage of the appraisal or failing to publish cost-effectiveness models), I identified the following themes as the leading Ground 2, substantive areas of appeal:

1. Failure to adequately attend to the ‘patient perspective’
2. Failure to identify patient subgroups for whom a drug may be cost-effective, even if it is not for the entire patient population
3. Failure to identify the right comparators
4. Failure to identify the right quality of life factors
5. Failure to identify a drug as ‘life extending’ or ‘innovative’ and thus subject to a higher cost-effectiveness threshold
6. Unreasonable management of uncertainty in cost effectiveness models
7. Unreasonable reliance on RCTs

While the way in which appellants claim that the evidence was unreasonably marshalled within each of these areas deserves future attention, for the focus of this chapter I draw on examples from across the seven that best illustrate the challenges in interpreting and weighing evidence.

⁴²⁴ Within this table, when indicated that Ground 1 was upheld 22 times, it means that in 22 appeals a Ground 1 appeal point was upheld. But, because multiple appellants can bring multiple Ground 1 appeal points, other Ground 1 appeal points may have been dismissed. When it indicates that Ground 1 was rejected on scrutiny 7 times, it means that NICE refused to allow one of the appellants to bring any Ground 1 appeal points to the Panel in 7 appeals.

It is worth noting, in backing up this chapter's earlier claim that patient involvement must be designed to elicit evidence rather than anecdote and emotional appeal, that non-specific claims that the patient perspective had not been adequately taken into account were sympathetically received but firmly dismissed. In the appeal hearing for the appraisal of cetuximab for head and neck cancer (TA145), the Mouth Cancer Foundation presented four of its members who "provided personal accounts of the suffering caused by head and neck cancer and of the very unpleasant nature of their treatments," by way of appealing that the Foundation's submission was not accorded the appropriate attention.⁴²⁵ The only response the Committee can give to this kind of appeal, given the public values it must consider, is to reiterate that it is fully aware of the seriousness of the condition and the desire of patients for the treatment in question. Indeed, the Committee members at the appeal defended themselves on this point by noting that they "clearly understood that they were assessing a treatment for a grave disease but that, however grave the disease, the treatment could only be recommended if there were good evidence that it was effective."⁴²⁶

Similarly, in TA72, the Arthritis and Musculoskeletal Alliance appealed on Ground 2 that:

...the Appraisal Committee had failed to give appropriate consideration to the patient perspective. In particular, the degree of clinical need of patients with rheumatoid arthritis did not appear to have been appropriately balanced against the wider financial implications."⁴²⁷

⁴²⁵ National Institute for Health and Care Excellence, "Decision of the Appeal Panel: Advice on Cetuximab for the Treatment of Locally Advanced Squamous Cell Cancer of the Head and Neck," August 22, 2007, <https://www.nice.org.uk/guidance/ta145/documents/head-and-neck-cancer-cetuximab-result-of-appeal-2>, 7.

⁴²⁶ Ibid., 8.

⁴²⁷ National Institute for Health and Care Excellence, "Decision of the Appeal Panel: Appraisal on the Use of Anakinra for the Treatment of Rheumatoid Arthritis," September 18, 2003, <https://www.nice.org.uk/guidance/ta72/documents/appraisal-of-the-use-of-anakinra-for-the-treatment-of-rheumatoid-arthritis-decision-of-the-appeal-panel2>, 14.

This claim, in the absence of any further evidence about how cost effectiveness considerations had unreasonably determined the decision, was swiftly rejected. The Appraisal Committee was able to point to the section of the Final Appraisal Determination that outlined their understanding of rheumatoid arthritis as proof that they had reasonably considered the patient perspective.

As these cases suggest, in an evidence-based approach that must consider the needs and values of the whole population, emotional appeals for reconsideration and basic claims that the patient perspective has not been taken into account appropriately will consistently play out like an attempt to square the circle. As argued in the previous section, the only way for Appraisal Committees to justifiably incorporate patient input and the only realistic pathway for patient groups to have an impact is if these groups can provide tangible evidence on the basis of unique patient and carer knowledge of patient experiences. On the one hand, this means that fostering deliberative resistance requires, far before the appeals process, that the patient involvement methods for technology appraisals be better designed with this in mind, that they provide *viable* processes for patients to contribute to decisions. On the other hand, this means that it is justifiable and indeed desirable that the initial scrutiny stage of the appeal process introduced in 2005 now adequately screens out such appeals that, despite good intentions, are not valid and will ring hollow in the context of a technology appraisal. Explaining to patient groups in the scrutiny letters *why* such points of appeal are invalid provides them with the knowledge necessary to bring the right kinds of arguments and evidence to appraisals and appeals in the future.

I continue now to cases in which all parties focused on the right kinds of claims but disagreed about whether evidence was admissible and, even if it was, what a reasonable interpretation of it entailed. If we ask patient groups to provide evidence

rather than anecdote, a key concern becomes the admissibility of the kind of ‘grassroots’ evidence that is collected and presented by patient interest groups to challenge assumptions made in the clinical or economic data. If money and time were not such determinative constraints on the appraisal process, it would be desirable for appropriately tailored, independent, representative, and robust social scientific studies to be carried out to flesh out the important gaps and uncertainties in the clinical and cost effectiveness data. In some cases, where cost-effectiveness is extremely sensitive to quality of life measurements that are clearly flawed in the clinical research and where time permits it, NICE may have a duty to fund such social science research. Given that robust studies will often not be possible, however, how should Appraisal Committees respond to surveys and similar on-the-fly data sets produced by patient groups?

In 2011, the pharmaceutical company Novartis and the patient group Kidney Cancer UK both filed appeals against a negative recommendation of everolimus for the second-line treatment of advanced renal cell carcinoma (TA219).⁴²⁸ Both brought Ground 2 appeals that the Appraisal Committee had unreasonably relied upon an average survival time for patients receiving best supportive care that was significantly longer than that seen in practice. This, the appellants claimed, resulted in a reduced incremental overall survival gain of everolimus compared to best supportive care. Because, on the Committee’s own admission, the survival gain was such a key factor in calculating the incremental cost effectiveness ratio, even slightly overestimating the survival time for best supportive care could serve to push everolimus beyond the cost effectiveness threshold. The Evidence Review Group (ERG), the expert external group that reviews evidence and reports on its strength to the Committee, had estimated the contested

⁴²⁸ National Institute for Health and Care Excellence, “Decision of the Appeal Panel: Advice on Everolimus for the Second-Line Treatment of Advanced Renal Cell Carcinoma,” February 28, 2011, <https://www.nice.org.uk/guidance/TA219/documents/renal-cell-carcinoma-second-line-metastatic-everolimus-appeal-decision2>.

survival time on best supportive care. The Committee conceded that “the ERG’s estimate of overall survival for patients receiving best supportive care...was higher than observed in clinical practice,” but it reasoned that the incremental *difference* in overall survival calculated by the ERG “was more plausible than that derived by the manufacturer and was based on all of the available data.”⁴²⁹

To attempt to address this uncertainty about average survival on best supportive care in clinical practice, Novartis had conducted an online survey of clinicians experienced in treating renal cell carcinoma, as well as reaching out to four UK hospitals for primary patient data on this question, which it received and analyzed.⁴³⁰ Both sources of clinical evidence suggested a lower average overall survival time that would alter the cost-effectiveness. However, the Appraisal Committee concluded that the data were based on a small sample and biased.⁴³¹ In the hearing, the Appeal Panel decided that the Committee had taken the survey “more seriously” than it had initially implied and “had considered all the relevant material and had considered overall survival in the round while giving most weight to statistical valid considerations.”⁴³² The Panel, the decision noted, “reminded itself that it should intervene *only if the guidance cannot reasonably be justified*,” and it found that this greater reliance on the statistical considerations of the incremental overall survival gain was justified (emphasis added).⁴³³ The appeal was therefore dismissed.

⁴²⁹ National Institute for Health and Care Excellence, “Final Appraisal Determination: Everolimus for the Second-Line Treatment of Advanced Renal Cell Carcinoma,” June 2010, <https://www.nice.org.uk/guidance/ta219/documents/renal-cell-carcinoma-second-line-metastatic-everolimus-final-appraisal-determination3>, 20.

⁴³⁰ National Institute for Health and Care Excellence, “Comments from Novartis Pharmaceuticals UK Limited on the Appraisal Consultation Document (ACD) for the Health Technology Appraisal of Everolimus for the Second-Line Treatment of Metastatic Renal Cell Carcinoma,” March 2010, <https://www.nice.org.uk/guidance/ta219/documents/manufacture-response-to-the-acd2>, 3-7.

⁴³¹ National Institute for Health and Care Excellence, “Decision of the Appeal Panel: Advice on Everolimus for the Second-Line Treatment of Advanced Renal Cell Carcinoma,” 11.

⁴³² *Ibid.*, 12.

⁴³³ *Ibid.*

This is an excellent example of where a reciprocal approach to public accountability will occasionally require more than showing that the outcome of a decision is ‘reasonable.’ Indeed, in this case, the Appeal Panel seems right in noting that the Appraisal Committee’s consideration of the evidence was reasonable. However, the Panel has elsewhere recognized that “it is possible to reach different conclusions, even on the same facts, without necessarily acting perversely.”⁴³⁴ And it seems a fair-minded person attempting to reason on the basis of mutually justifiable reasons could also have given the evidence provided by Novartis more weight, in light of the admitted shortcomings of the initial survival estimate. This could have produced a different, reasonable conclusion. Perhaps the Committee finds, as it may in cases of such deliberative disagreement, that on reflection and – if necessary – majority vote it simply perceives one weighing up of the evidence to be *more* reasonable. This process of handling the evidence would meet the demands of Daniels and Sabin’s accountability for reasonableness.

However, it is not sufficient on the conditions of my framework. To foster resistance to such *reasonable* disagreement about the weighing of evidence, NICE has a duty to provide an account of how Novartis (or patient groups) could in the future leverage such patient data: what more, in other words, would Novartis have had to do to render its evidence reliable in the appraisal context? If the sample size was too small or biased, what would be an appropriate sample size or study design, given limitations based on the total patient population size and the available patient data? By sketching these answers, we can recognize the legitimacy and reasonableness of the Committee’s decision while also understanding that reciprocity requires a particular type of response –

⁴³⁴ National Institute for Health and Care Excellence, “Decision of the Appeal Panel: Appraisal of the Use of Photodynamic Therapy for Age-Related Macular Degeneration,” March 17, 2003, <https://www.nice.org.uk/guidance/TA68/documents/decision-of-the-appeal-panel-appraisal-of-the-use-of-photodynamic-therapy-for-agerelated-macular-degeneration-2>, 7.

institutions that support contestation – in the face of deliberative disagreement. It is likely that the Appeals Panel is not the most well positioned component of NICE to provide this response; that is more likely to be the original Appraisal Committee or Evidence Review Group. Nevertheless, the Panel *is* the most well positioned group to identify when this kind of response needs to be given. In addition to addressing the appellants in a given case, such responses should also feed back to the Public Involvement Programme that provides the evidence submission forms and guidance for patient groups. These need to be adjusted to give an account of the *standards* for collection and analysis of evidence that patient groups are expected to meet.

The use of these sorts of grassroots surveys and polls is not limited to manufacturers, although their significantly greater resource levels allow for more extensive and well-designed data collection. In the appeal of TA85 on immunosuppressive agents for renal transplantation, the patient interest group National Kidney Federation UK carried out a “straw poll” of its members that found that 50% were taking a drug that the Appraisal Committee restricted the use of; this data collection was part of a larger appeal point about the expected negative effects of the recommendation on patients.⁴³⁵ At the appeal hearing, one of the attending Appraisal Committee members responded that he “doubted that 50% of post renal transplantation patients were taking” the drug.⁴³⁶ This may be a reasonable response; in fact, to the extent that this straw poll lacked any other context or design, a fair-minded person would be less inclined to think it meets relevant evidence standards than the Novartis research. Nevertheless, it is crucial that NICE take such opportunities to clarify how exactly

⁴³⁵ National Institute for Health and Care Excellence, “Decision of the Appeal Panel: Appraisal of the Use of Immunosuppressive Therapy for Renal Transplantation,” April 26, 2004, <https://www.nice.org.uk/guidance/ta85/documents/appraisal-of-the-use-of-immunosuppressive-therapy-for-renal-transplantation-decision-of-the-panel-2>, 20.

⁴³⁶ Ibid.

patient groups can leverage their connection to the patient base to provide Appraisal Committees with evidence from patients that may actually be taken seriously. If a ‘straw poll’ is not a method that produces convincing evidence, what method would NICE recognize as providing reliable evidence on the accurate landscape of patient experience?

Beyond quantitative evidence, patient groups often gather and provide the Appraisal Committee with large amounts of qualitative evidence, particularly when quality of life considerations are contested. In one of the most high-profile appraisal cases, TA111 for Alzheimer’s drugs, NICE found that treatment was cost effective in the moderate stage of the disease, but not the mild stage. The Appraisal Committee was flooded with responses that echoed the following from a clinician:

I would ask all members of the committee to consider if they had dementia and could have treatment that would benefit them cognitively and effectively stabilize their condition for between one and two years, would they want that at the stage of dementia when they still had insight, significant retention of activities of daily living and reasonable social function, or would they rather wait until they had deteriorated to a point where they were becoming behaviourally disturbed, were losing insight and at a point where they were already needing substantial ‘hands on’ care and supervision.⁴³⁷

The leading patient and carer interest group, the Alzheimer’s Society, replied on the basis of similar objections from their members, stating, “We know from our membership that people with dementia and their carers value stabilization at the earliest point, as this enables the best possible quality of life to be maintained for as long as possible.”⁴³⁸ During the Appeal hearing, the Committee did indeed “acknowledg[e] that they had

⁴³⁷ National Institute for Health and Care Excellence, “Drugs for Alzheimer’s Disease: NICE Response to Consultee and Commentator Comments on the ACD-2 (January 2006),” May 2006, <https://www.nice.org.uk/guidance/ta111/documents/alzheimers-disease-responses-to-consultee-and-commentator-comments2>, 4.

⁴³⁸ Alzheimer’s Society, “Alzheimer’s Society Response to NICE Appraisal Consultation Document 2,” February 2006, <https://www.nice.org.uk/guidance/ta111/documents/alzheimers-consultee-and-commentator-comments-alzheimers-society2>, 2.

heard anecdotally that improvement in the mild phase of the disease was more meaningful.”⁴³⁹

Yet the appeal decision noted:

The Appraisal Committee Chairman addressed the assertion...that treatment was more meaningful in early dementia. The Committee’s role was to assess the evidence. The Panel heard that there was no evidence of a differential effect in respect of the impact of treatment, and evidence of no differential effect, as described in Technical Report 1.⁴⁴⁰

In other words, the Committee was unable to locate an evidence base, beyond the anecdotal submissions, for this ‘meaningfulness’ in using the trial data and statistical analysis. Particularly because of the plausibility of the argument that treatment in dementia is more meaningful in the early rather than later stages of the disease, and because of how traditional quality of life measures might struggle to capture this feature, it seems that fair-minded people could have come to a different but reasonable conclusion in this case. Again, this is true even if the judgement the Appraisal Committee made is also reasonable. And, in such a case of deliberative disagreement, after all have attempted to reason on the basis of relevant standards and values, the majority opinion of the Committee should be allowed to stand. But NICE should recognize in the aftermath the duty to explain to the Alzheimer’s Society what standards of evidence must be met if the Society is to draw on its membership to show that “people with dementia and their carers value stabilization at the earliest point.”

What I hope these examples suggest is the existence of deliberative disagreement about the interpretation of evidence in technology appraisals. This disagreement is

⁴³⁹ National Institute for Health and Care Excellence, “Decision of the Appeal Panel: Donepezil, Rivastigmine, Galantamine and Memantine for the Treatment of Alzheimer’s Disease, Appeal by the Alzheimer’s Society, Age Concern, Counsel and Care, the Dementia Care Trust and the Royal College of Nursing,” July 14, 2006, <http://www.nice.org.uk/guidance/TA111/documents/appeal-by-the-alzheimers-society-age-concern-counsel-and-care-the-dementia-care-trust-and-the-royal-college-of-nursing2>, 5.

⁴⁴⁰ Ibid.

certainly not grounds, on my framework of public accountability, to doubt the legitimacy of a decision that is the result of process in which all tried to reason reciprocally. But in recognizing that others could have reasonably interpreted the evidence otherwise, we must be willing to support avenues for the right kind of contestation in the future. This means both adjusting initial patient involvement in appraisals so that it is tailored to evidence provision (and thus capable of having real influence within the process), and ensuring that an appeals process includes flagging up the need for Evidence Review Groups or the Appraisal Committee to provide guidance on how evidence from patients can be collected in the future in a way that is recognizably reliable.

Particularly given the number of different ‘steps’ that I have taken within this chapter, it is useful to summarize the arguments before making some concluding remarks about the chapter’s unique role as a case study of public accountability. In contrast to the HFEA, NICE has already committed to a reciprocal method of reasoning, that of accountability for reasonableness. While many of the arguments I made in the HFEA chapter about member appointments and the goals of public involvement may indeed be applied to refine the current state of the NICE Board or the Citizens Council and their identification of ‘relevant’ principles in *Social Value Judgements*, I recognize these as being broadly in line with my framework even if there may be room for improvement. Instead, I focused on the technology appraisal process, understanding that even when an arm’s length body has established a fair process and identified the public reasons upon which decisions should be based, it still must substantiate this set of values in light of the arguments and evidence given within particular cases.

Because the values of technology appraisal are properly set in advance by processes accountable to all citizens, patients and patient groups involved in appraisals

cannot dictate which norms should guide the Appraisal Committee or which value is the overriding one in a specific case. The only justifiable influence they can have on the Committee's reasoning is through evidence that they may uniquely provide about patient experiences and which may affect the principles of clinical effectiveness and cost effectiveness. The patient voice may function as evidence in this way by filling holes in RCT data or actively shedding light on flawed assumptions in cost-effectiveness models about quality of life factors or clinical realities. NICE must reconstruct its patient involvement processes in recognition of this justified role for the patient voice. In particular, I suggested that the 'patient expert' is a poor method, and that NICE should rather emphasize evidence submissions from patient groups, who may then exert their organizational resources towards the collection of representative information from patient-members. This adjustment will both provide NICE with better evidence about relevant considerations in appraisals and ensure that the patient voice does not go unheard or unheeded.

However, once we resolve this initial confusion around the role of patients by casting them as evidence providers, a second practical challenge for Appraisal Committees comes into sharp relief: how can we evaluate the strength of this evidence and weigh it against that from RCTs or cost effectiveness models in a publicly accountable manner? We may not be able to agree on standards of validity for such evidence and, even when we do, the burdens of judgment suggest that reasonable people may nevertheless disagree about what evidence entails. Through a review of the appeals process, I showed how NICE currently, in keeping with Daniels' and Sabin's understanding of the demands of accountability for reasonableness, sees its duty met when it confirms that an Appraisal Committee's decision was not an *unreasonable* interpretation of the evidence. However, the conception I have endorsed of

accountability as reciprocity requires taking seriously the fact that other reasonable interpretations may exist that support incompatible but reasonable conclusions. This requires an adjustment to what is otherwise a well constructed and deliberative appeals process. When the Appeal Panel notes, as it often does in dismissing Ground 2 appeals, that the decision of the Appraisal Committee was a reasonable interpretation of the evidence, it should be required to refer the Evidence Review Group⁴⁴¹ and the Public Involvement Programme back to those points of appeal in which patients' evidence was found to be relevant but not robust or reliable enough to weigh up against technical reports in the eyes of the Appraisal Committee. This will enable NICE to meet the condition of fostering the means of deliberative resistance in two ways. First, it provides the necessary explanation for the specific appellant in question to shape their future involvement in appraisals and appeals. Second, by institutionalizing a feedback loop from the appeals process (the end of an appraisal) to the Public Involvement Programme (the beginning), it may improve the general guidance on evidence submission, thus supporting the best avenue for all future patients and groups to engage with the appraisal process.

In addition to these fairly NICE-specific recommendations, this chapter provides two unique contributions as a case study of substantiating public accountability in the arm's length context. First, it develops the idea of fostering deliberative resistance and the kinds of procedural reforms that follow from this fourth condition of my framework. This is significant given that it is not a condition of accountability that is properly attended to in other practical applications of public reason and reciprocity, like Daniels' and Sabin's work. Neither is it a concept particularly well developed in the theoretical

⁴⁴¹ In single technology appraisals (STAs), it is an Evidence Review Group that reviews the evidence, but in multiple technology appraisals (MTAs), it is an Assessment Group.

literature, although the work of Jane Mansbridge – from whom the phrase is borrowed – is a notable exception.⁴⁴² Second, this NICE case study illustrates how the implementation of a theoretical framework cannot be expected to *simply* resolve practical challenges. Indeed, as my development of patient-involvement-as-evidence-provision shows, a framework of public accountability based upon reciprocity will likely have knock-on effects: it may be expected to present new practical challenges even as it resolves others. However, by reference to the way in which the appeals process should manage reasonable disagreement about evidence, and how this should feed back to the design of patient involvement, I have tried to show how a reiterative application of my four conditions can continue to refine the problems that arise, ensuring that processes are legitimate.

⁴⁴² Mansbridge, “Using Power/Fighting Power: The Polity.”

Chapter VI. Regulatory Boundaries in the MHRA: Opening up the Concepts of Safety and Effectiveness

Institutionally, the MHRA may initially appear out of place next to the HFEA and NICE. Indeed, this policy-making setting exhibits a number of different characteristics from the previous two. Although it has undergone a variety of nominal transformations since 2003, such as merging the medicines and medical device regulatory agencies, its primary function of evaluating the safety and effectiveness of medical products dates back to the 1960s, making it significantly older and more well-established than the other two. Structurally, it is set up as an executive agency rather than a non-departmental public body, meaning that it technically has more direct lines of accountability to the Department of Health and is often overlooked in the ‘quango’ debate. Furthermore, its structure has a supranational dimension, as the European Commission hands down a variety of medicines and medical devices regulations that the Agency is responsible for implementing within the UK. There has been relatively little attention paid to the MHRA in the field of bioethics or at the intersection of ethics and public policy, whereas the HFEA and NICE have garnered a great deal of attention in both areas. Perhaps most importantly, unlike NICE and the HFEA, the MHRA is not immediately recognizable as a body that ‘does’ bioethics: both external observers and the MHRA itself do not tend to see the institution as making the type of value-based decisions on behalf of society that the HFEA and NICE undertake.

Rather, the decisions that it makes about what drugs and devices should be licensed are most often represented as purely fact-based, expert judgements. The MHRA stresses that “[a]ll [its] work is underpinned by robust and fact-based judgements to ensure that the benefits justify any risks,”⁴⁴³ and that its “regulatory decisions are impartial and based solely on the extensive evidence of quality, safety, and efficacy required for each product.”⁴⁴⁴ Most members of the public have neither the time nor relevant knowledge and skills necessary to sift through evidence from clinical trials and pharmacovigilance reports in making choices about what medicines to use, the argument goes, so we rationally defer these decisions to respected scientists and clinicians, trusting them to appropriately measure and balance the risks and benefits of medical products that are allowed on the market. On this understanding, the MHRA is the paradigm of a government agency set up at arm’s length to draw upon technical expertise, one of the three justifiable reasons enshrined in the Maude tests for removing a decision from democratically-elected officials.

As a democratic society, we are generally more comfortable with this kind of delegation to technocrats because it is not understood as a deferral of authority in identifying the public or social *values* that should guide regulatory decisions, which – in contrast – we do recognize to be the case for the decisions that we abdicate to the HFEA and NICE. Rather, guiding values are imparted to the MHRA by the Department of Health and almost universally endorsed by the public: the protection and promotion of public health based on consideration of whether new medical products are safe and whether they work. The MHRA’s role is supposed to be ‘limited’ to drawing on

⁴⁴³ See, for example, “Background Information,” in Medicines and Healthcare Products Regulatory Agency, “GVK Biosciences: MHRA Responds to European Medicines Agency’s Recommendation to Suspend Medicines,” *Press Release*, January 23, 2015, <https://www.gov.uk/government/news/gvk-biosciences-mhra-responds-to-european-medicines-agencys-recommendation-to-suspend-medicines>.

⁴⁴⁴ Medicines and Healthcare Products Regulatory Agency, “Medicines and Medical Devices Regulation: What You Need to Know,” 4.

technical expertise to meet these pre-ordained public values. Its current model of public accountability, in turn, is based upon its ability to wield this expertise in meeting those values. It demonstrates its accountability to the public in spite of its independent decision-making and rarefied technical deliberations to the extent that the medical products that members of the public use are, in fact, safe and effective.

On the whole, this model of accountability for the type of independent decision-making that the MHRA undertakes works well. We value this division of epistemic labour that means that most of us, most of the time, do not need to think about whether aspirin bought over the counter or amoxicillin prescribed by a GP will have serious or unforeseen side effects, or whether an inhaler valve will deliver the proper dose of medicine in the advent of an asthma attack. Although serious failures are highly publicized and often lead to review and reform, such as the recent PIP breast implant scandal in which it transpired that implants imported from a French manufacturer used industrial rather than medical grade silicone,⁴⁴⁵ generally the public trusts that the experts at the MHRA are capable of accessing and interpreting evidence on medicines and devices and making regulatory decisions that promote public health on behalf of society. Indeed, in a representative sample of UK adults, nearly 9 out of 10 believe medicines and medical devices are safe to use.⁴⁴⁶ Even recent works that have directed popular attention to the larger, structural issues of medical research and drug regulation in the UK, such as Ben Goldacre's pop-sci bestseller *Bad Pharma*, do not question this technocratic model of regulation. If anything, concerns like Goldacre's about regulatory capture by the pharmaceutical industry or the poor and incomplete data coming out of

⁴⁴⁵ Carl Heneghan, "The Saga of Poly Implant Prosthese Breast Implants," *BMJ* 344, no. 2 (2012).

⁴⁴⁶ Ipsos MORI, "Risks and Benefits of Medicines and Medical Devices – Perceptions, Communication and Regulation: General Public Quantitative Survey," Report for the MHRA, (2009), <http://www.mhra.gov.uk/home/groups/comms-ic/documents/websiteresources/con052027.pdf>.

clinical trials suggest that the MHRA's independence and technical expertise need to be bolstered to combat harmful incentives within the pharmaceutical development process.

However, emerging biotechnologies – in particular, new devices – are exposing holes in this account of the limited, technical function of the MHRA. In Chapter II, I laid out the practical challenge for the MHRA of regulating so-called 'borderline' products, drawing on the recent Stephenson report and its sweeping recommendations for the reform of the MHRA's devices division, as well as concurrent reform efforts at the European level.⁴⁴⁷ This challenge posed by novel borderline products, I suggested, can be understood as a manifestation of the legitimacy problem because there is real uncertainty in such borderline cases about what protecting and promoting patient safety and health requires, which threatens the MHRA's method of accounting to the public by ensuring the safety and effectiveness of such products. Direct-to-consumer genetic tests that offer health information, telehealth systems that turn the home into a point of health and social care, health and lifestyle 'apps' purchased on phones and tablets, and devices promising cognitive enhancement have all confounded traditional regulatory pathways for and methods of evaluating safety and effectiveness. Although each type of technology raises its own set of challenges for the regulator, they all similarly resist the MHRA's standard risk-based classification and risk-benefit evaluation.

The Stephenson report, then, has usefully identified these borderline cases as critical threats to the MHRA's public accountability, given the inability of current regulatory pathways and methods to find purchase in analyzing these promising new technologies, and the resultant worry that patients will be exposed to unsafe or

⁴⁴⁷ Stephenson, "Expert Clinical Advice - MHRA Medical Devices"; European Commission, "Executive Summary of the Impact Assessment on the Revision of the Regulatory Framework for Medical Devices."

ineffective products. Indeed, Professor Stephenson stressed in the introduction to the report:

We need a medical devices regulatory system which encourages the devices industry to develop new technologies which improve the quality of our lives. However, the same system needs to predict risk, detect when harm is occurring and be capable of intervening swiftly to limit adverse events.⁴⁴⁸

The response to the borderline challenges introduced by these new technologies has been the attempt to reconsolidate the authority and sufficiency of technical expertise in such regulatory decision-making. The Stephenson report's full (and telling) title is *Expert Clinical Advice – MHRA Medical Devices: Report of the independent review on MHRA access to clinical advice and engagement with the clinical community in relation to medical devices*. Indeed, the key recommendations of the report are focused on the “organisation of clinical advice input, resources and leadership” in the devices division: creating a formal mechanism for incorporating clinical advice, hiring clinically trained staff within the MHRA, and developing an external network of clinical advisors.⁴⁴⁹ In accepting these recommendations, the primary adjustment within the MHRA has been the creation of the new Devices Expert Advisory Committee, which will “help the agency have stronger links with the wider scientific community to facilitate access to specialist expertise,”⁴⁵⁰ which will in turn “ensure that medical devices are acceptably safe and are used both safely and effectively.”⁴⁵¹

⁴⁴⁸ Stephenson, “Expert Clinical Advice - MHRA Medical Devices,” 4.

⁴⁴⁹ *Ibid.*, 5.

⁴⁵⁰ Medicines and Healthcare Products Regulatory Agency, “New Independent Expert Advisory Group to Be Created as the MHRA Responds to Stephenson Review,” Press Release, (July 18, 2014), <http://webarchive.nationalarchives.gov.uk/20141205150130/http://www.mhra.gov.uk/NewsCentre/Pressreleases/CON437435>.

⁴⁵¹ Medicines and Healthcare Products Regulatory Agency, “New Chair Appointed for the Newly-Formed Devices Expert Advisory Committee,” Press Release, (February 18, 2015), <https://www.gov.uk/government/news/new-chair-appointed-for-the-newly-formed-devices-expert-advisory-committee>.

In this chapter, I suggest how the tactic of bolstering clinical expertise has fundamentally misunderstood the basic problem posed by borderline products. It may be the case, as the Stephenson Committee suggested, that the MHRA's clinical expertise on medical devices has lagged behind that on medicines: it is certainly true that the methods for clinically assessing the two differ substantially, as the safety of devices tends to depend much more upon physical design and the way in which they are used than the 'active ingredients,' a main consideration for medicines.⁴⁵² But this is not the primary problem with the MHRA's approach to public accountability that is exposed by new 'borderline' products like cognitive enhancement devices or telehealth systems. Rather, these new technologies starkly expose the fact that what it *means* for a device to be safe and effective are value judgements and ones which may invoke a much broader ranges of reasonable considerations than those traditionally identified, calculated, and managed by scientific experts.

We may agree – and I take as given in this chapter – that the broad considerations of safety and efficacy in deciding what medical products should be licensed are the properly 'public' or 'mutually justifiable' ones - that these two public values present legitimate substantive principles for making licensing decisions at arm's length. In this sense, it seems that Parliament has publicly "set" the normative standards and merely asks the MHRA to "meet" them, as suggested by the discussion in Chapter III about standard-setting versus standard-meeting bodies. But what it means in any given case, as consideration of borderline products will illustrate, for a product to 'work' and what constitutes the 'risks' and 'benefits' they present are not questions that can or should be settled *a priori* by clinical experts. On the one hand, this means that the experts may not

⁴⁵² Peter Wilmshurst, "The Regulation of Medical Devices," *BMJ* 342 (2011); Rod S. Taylor and Cynthia P. Iglesias, "Assessing the Clinical and Cost-Effectiveness of Medical Devices and Drugs: Are They That Different?," *Value in Health* 12, no. 4 (2009); Deborah Cohen, "Out of Joint: The Story of the ASR," *BMJ* 342 (2011).

always know what factors are relevant in the calculation of risks and benefits. And, on the other hand, this means that the value placed on different types of benefits, and thus the willingness to accept risk in pursuit of those benefits, may vary greatly between reasonable individuals. As such, on the way to meeting the standards of safety and effectiveness, the MHRA must confront and settle a second-order series of value judgements. In making these judgements, about what factors count as risks and benefits and what represents a reasonable risk-benefit profile, the MHRA performs a role far more analogous to that of the HFEA and NICE: it must make value judgements on behalf of society at arm's length that result in coercive decisions about which novel biotechnologies citizens can access. In making such decisions, reciprocity requires that these judgements be made on the basis of mutually acceptable reasons. As such, the MHRA must open up these judgements to deliberative scrutiny and adjust its procedural and structural requirements to facilitate the search for reciprocal reasons about the weighing of risks, benefits, and effectiveness in deciding whether products should be licensed or banned.

This means that the MHRA must relinquish the notion that bolstering clinical expertise can solve the problem of ensuring the safety and effectiveness of novel 'borderline' biotechnologies, because these concepts are not settled in advance such that clinical experts can simply cash out a risk-benefit analysis on the basis of their scientific expertise and existing evidence. Instead, the interpretation of these values, of whether something is safe and works, must also draw on the values and perspectives of the public and patients on whose behalf such judgements are made, as in the HFEA's public consultation process, NICE's Citizens Council, and NICE's patient involvement in appraisals. Returning to Professor Stephenson's account above, public accountability demands that when an arm's length body like the MHRA makes collective decisions to

restrict citizens from accessing new medical products, reciprocal reasoning about public values – not the values asserted by technical authorities – must determine what it means for a product to “improve the quality of our lives” and “when harm is occurring.”⁴⁵³

In the preceding chapter on NICE, I began by endorsing its approach to accountability for reasonableness and taking as given its current ‘relevant’ values outlined in *Social Value Judgements* in order to illustrate how we can substantiate the practical requirements of this kind of general framework for the process of patient involvement in technology appraisals. Here, I similarly accept from the outset that the MHRA focuses on the right broad considerations of risk, benefit, and effectiveness, in order to think about how we can substantiate these general values in a publicly accountable way within the context of particular licensing decisions. I take it that no reasonable conception of drug and device licensing can reject the relevance of at least some standards of safety and effectiveness. Libertarian arguments about the right to purchase any manufactured product on an unfettered free market, however dangerous or ineffective, are checked by the interests of other citizens in a licensing system that assures their health and safety – and that of those who are not fully autonomous, like children – given limited time, resources, and knowledge. In this sense, the case study of the MHRA is of a similar type to that of NICE. I draw some parallel conclusions about the identification of quality of life factors, public and patient involvement processes, and the type of evidence that must be brought to bear on decision-making. However, because the borderline cases of cognitive enhancement and telehealth systems that I address expose the basic fluidity of some of the fundamental regulatory distinctions upon which the MHRA relies, I also argue for certain adjustments to its regulatory boundaries, which

⁴⁵³ Stephenson, “Expert Clinical Advice - MHRA Medical Devices,” 4.

usefully differentiates the contribution of this case study chapter from that of the previous in demonstrating what follows from the application of my framework.

The chapter proceeds as follows: I set out the MHRA's most basic regulatory distinction between medical and non-medical/consumer products, one of the three 'blurred lines' discussed in Chapter II. In the following section, I suggest four common rationales for this distinction between the 'medical' and the 'non-medical.' In section three, I argue that the borderline problems caused by transcranial direct current stimulation (tDCS) devices, which may be used for consumer enhancement or medical treatment, stem from its inability to map onto this traditional distinction. This highlights the need to consider risks and benefits in regulating tDCS for enhancement despite its 'non-medical' nature, but simultaneously illuminates the importance of considering benefits beyond 'medical' benefits in regulating tDCS for cognitive enhancement. Section four uses the borderline problems of telehealth systems and the home context to suggest analogously the need to consider risks beyond 'health' risks. Section five concludes with how the processes and regulatory boundaries of the MHRA must be adjusted on the basis of these findings in order to make decisions in a publicly accountable manner.

1. When is a product a 'medical' product?

The MHRA states, "As a general rule, products making medical claims will be regulated either by the medical devices regulations or by medicines legislation."⁴⁵⁴ As such, in determining whether a product is medical or non-medical and thus whether the

⁴⁵⁴ Medicines and Healthcare Products Regulatory Agency, "Borderlines between Medical Devices and Medicinal Products," Guidance on legislation, (June 2013), <https://www.gov.uk/government/publications/borderlines-between-medical-devices-and-medicinal-products>, 3.

MHRA is responsible for assessing it, a key factor is whether it makes any medical claims. What, then, does it mean to make a ‘claim’ and when are claims ‘medical’ in nature? There are some confusing discrepancies between the medical device and medicines regulatory regimes in determining the answer to this question. The current definition⁴⁵⁵ of ‘medical device’ is:

Any instrument, apparatus, appliance, software, material or other article, whether used alone or in combination, including the software intended by its manufacturer to be used specifically for diagnosis and/or therapeutic purposes and necessary for its proper application, *intended by the manufacturer* to be used for human beings for the purpose of:

- a) diagnosis, prevention, monitoring, treatment or alleviation of disease,
- b) diagnosis, monitoring, treatment, alleviation of or compensation for an injury or handicap,
- c) investigation, replacement or modification of the anatomy or of a physiological process,
- d) control of contraception;

and which does not achieve its principal intended action in or on the human body by pharmacological, immunological or metabolic means, but which may be assisted in its function by such means (emphasis added).⁴⁵⁶

A product is identified as a medical device, according to this definition, based both on its *intended purpose* (for diagnosis and/or therapeutic purposes, to distinguish it from non-medical devices) and the means by which it *functions* (not primarily through pharmacological, immunological, or metabolic means, to distinguish it from medicines). In ascertaining whether a product’s intended purpose is medical, the definition suggests that the relevant perspective is that of the *manufacturer*.

⁴⁵⁵ Active implantable medical devices and *in vitro* diagnostic medical devices have more specific regulations under two separate directives, but I will consider them within the broad class of medical devices in this chapter as principles for their regulation are similar.

⁴⁵⁶ Medicines and Healthcare Products Regulatory Agency, “Borderlines between Medical Devices and Medicinal Products,” 4.

In contrast, the definition of a medicinal product (a medicine), makes no reference to the intent of the manufacturer:

- a) Any substance or combination of substances presented as having properties for treating or preventing disease in human beings; or
- b) Any substance or combination of substances which may be used in or administered to human beings either with a view to restoring, correcting or modifying physiological functions by exerting a pharmacological, immunological or metabolic action, or to making a medical diagnosis.⁴⁵⁷

The first half of this two-pronged definition means that a product may be defined as a medicine based on its presentation. And, as the MHRA explains, rather than relying on the perspective of the manufacturer in determining whether a product is “presented as having properties for treating or preventing disease,” the European Court of Justice has ruled that this determination must be based on the perspective of the ‘averagely well-informed consumer’ and the impression that he makes about whether the product is intended to treat or prevent disease.⁴⁵⁸

For instance, the MHRA notes:

A bath foam which bears a claim such as ‘relaxing’ or which mainly refers to the ‘well being’ will normally be considered by the consumer as a typical cosmetic product. On the other hand, such a bath foam may be considered to be a medicinal product if the product is mainly presented as a treatment for cold or flu.⁴⁵⁹

This illustrates the importance of how a product’s claim will “normally be considered by the consumer”: the premise of the bath foam example is that an averagely well-informed consumer will not consider a product to be medical simply if it claims to improve well-being or aid relaxation. Even products that make claims about health are not always

⁴⁵⁷ Ibid., 3-4.

⁴⁵⁸ Medicines and Healthcare Products Regulatory Agency, “A Guide to What Is a Medicinal Product,” MHRA Guidance Note No. 8, (November 2012), https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/398998/A_guide_to_what_is_a_medicinal_product.pdf, 9.

⁴⁵⁹ Medicines and Healthcare Products Regulatory Agency, “Borderlines between Medical Devices and Medicinal Products,” 5.

considered by the average consumer to be medical products. The MHRA advises that it “does not normally consider claims to ‘maintain’ or ‘help to maintain’ or ‘support’ health or a healthy lifestyle, as medicinal in themselves. Nor, if such claims are clearly made in relation to healthy bodily functions or organs is the MHRA likely to consider them as presenting the product for treating disease.”⁴⁶⁰ Only if a product’s claim suggests that it “may restore, or help to restore, a specific bodily function or organ to a normal healthy state,” will it be considered medical.⁴⁶¹ The European Commission similarly notes that although there is no Community-definition of disease to use in interpreting the limits of ‘treating and diagnosing disease,’ the European Court of Justice has ruled:

A product presented as counteracting certain conditions or sensations, such as heaviness in the legs or tiredness or itching is not per se a medicinal product... Since these sensations may have no *pathological significance*, ‘a reference to such states or sensations in the presentation of a product is not decisive’ (emphasis added).⁴⁶²

In other words, medical products must treat or diagnose *disease*, which in turn must have a *pathological* origin.

In terms of consistency in the MHRA’s regulatory approach, there is no sound reason why the intention of the manufacturer should determine the medical nature of a product in the area of devices but the perception of the consumer should determine the medical nature of a product in the area of ‘substances.’ In fact, in conflict with the formal definition, guidance for medical device regulation suggests the insufficiency of focusing solely on the intent of the manufacturer. Guidance on the status of medical software including apps notes, “General disclaimers (eg ‘this product is not a medical

⁴⁶⁰ Medicines and Healthcare Products Regulatory Agency, “A Guide to What Is a Medicinal Product,” 10.

⁴⁶¹ Ibid.

⁴⁶² Medicines and Healthcare Products Regulatory Agency, “Borderlines between Medical Devices and Medicinal Products,” 6.

device’) are not acceptable if medical claims are made or implied...Anecdotal quotes and testimonials are considered to be implied claims.”⁴⁶³ And, in the course of revising the medical devices directives, a European Commission staff working group found that “according to the prevailing interpretation of the Commission, Member States and stakeholders, a device falls within the definition of a medical device *when it pursues a medical purpose*” (emphasis added).⁴⁶⁴ This, in contrast to manufacturer’s intent or averagely well-informed consumer’s impression, suggests no subjective perspective from which a medical purpose is determined.

This relates to the second prong of the definition of a medicine, which suggests that regardless of a substance’s presentation and any explicit or implicit claims made about preventing or treating disease, it may be classed as a medicine if it functions to restore, correct or modify physiological functions through pharmacological, immunological or metabolic action. This prong removes the subjectively determined existence of medical claims (whether from the perspective of the consumer or the manufacturer) and determines medical purpose exclusively on the basis of function. Likewise, the definition of medical device includes those intended for the “investigation, replacement or modification of the anatomy or of a physiological process,” a clause which similarly focuses solely on function and does not require claims of diagnostic or therapeutic purpose. Both of these clauses, then, suggest that a product need not in fact make any diagnostic or therapeutic claims to be classified as medical: it must merely ‘modify’ physiological functions or anatomy.

⁴⁶³ Medicines and Healthcare Products Regulatory Agency, “Medical Device Stand-Alone Software Including Apps,” Guidance, (August 8, 2014), <https://www.gov.uk/government/publications/medical-devices-software-applications-apps/medical-device-stand-alone-software-including-apps#more-information>.

⁴⁶⁴ European Commission, “Commission Staff Working Document: Impact Assessment on the Revision of the Regulatory Framework for Medical Devices” (Brussels, 2012), <http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52012SC0273&from=EN>.

Both, however, are not in practice considered sufficient in distinguishing between medical and non-medical products. Considering the borderline between medicines and cosmetics, the European Commission observes that all moisturising creams modify physiological function through a metabolic action by adding water to skin cells, thus meeting the second prong definition of a medicinal product, but such creams are cosmetic rather than medical in nature.⁴⁶⁵ Citing the European Court of Justice decisions, the Commission notes, “[T]he modification has to be more than insignificant.”⁴⁶⁶ Furthermore, even substances that *do* modify physiological functions significantly, as evidenced by the dose-response relationship, are not always considered to be medicines: both alcoholic beverages and coffee or other caffeinated drinks, for instance, are regulated as foods despite presenting similar dose-response relationships to some medicines.⁴⁶⁷ With respect to medical devices, as Hannah Maslen and a group of academics at the Oxford Martin School point out, a solitary focus on a product’s “investigation, replacement or modification of the anatomy or of a physiological process” would mean implausibly including books, DVDs, and computer games in the category of medical devices given that they modify neuronal connections.⁴⁶⁸ The authors cite consensus within the European Community that having a diagnostic or therapeutic purpose is still requisite for a product to be classified as a medical device.⁴⁶⁹

⁴⁶⁵ European Commission, “Guidance Document on the Demarcation Between the Cosmetic Products Directive 76/768 and the Medicinal Products Directive 2001/83 as Agreed Between the Commission Services and the Competent Authorities of Member States” (Brussels, April 30, 2004), http://ec.europa.eu/growth/sectors/cosmetics/products/borderline-products/docs/guidance_doc_cosm-medicinal_en.pdf, 7.

⁴⁶⁶ *Ibid.*, 8.

⁴⁶⁷ Dirk W. Lachenmeier et al., “What Is a Food and What Is a Medicinal Product in the European Union? Use of the Benchmark Dose (BMD) Methodology to Define a Threshold for ‘pharmacological Action,’” *Regulatory Toxicology and Pharmacology* 64, no. 2 (2012).

⁴⁶⁸ Hannah Maslen et al., “The Regulation of Cognitive Enhancement Devices: Extending the Medical Model,” *Journal of Law and the Biosciences* 1, no. 1 (2014): 74.

⁴⁶⁹ *Ibid.*, 75.

In sum, despite the many-limbed and often inconsistent definitions of medicines and medical devices, some general conclusions may be drawn about the current boundary that the MHRA upholds between medical and non-medical products. Medical products serve a diagnostic or therapeutic purpose, where diagnosis or treatment is limited to disease, disability, or symptoms of pathological significance. To protect consumers from the opportunity cost of using products whose efficacy has not been assessed, a product whose presentation reasonably suggests or implies that it serves a diagnostic or therapeutic purpose is considered to be claiming a medical purpose and is subject to medical regulation. Armed with this understanding of how the MHRA delimits the ‘medical,’ I turn to the questions of why the MHRA works so hard to preserve this boundary and how it justifies special assessment of safety and effectiveness for products falling in the medical sphere.

2. Why is the medical sphere regulated uniquely?

Medical products – both medicines and medical devices – are subject to different, pre-emptive, and generally more robust regulation and oversight than seemingly similar products that fall outside the medical sphere, such as certain cosmetics, food, and consumer gadgets and software. Medical products are subject to *pre*-marketing assessment of risks and benefits while similar consumer products are not. Therefore, getting the medical/consumer classification ‘right’ is crucial because it determines the scrutiny of risks and benefits that a product will receive, as well as post-marketing follow-up to detect unexpected harms. Furthermore, medical products must demonstrate effectiveness, while consumer products need not. In *Law and the Regulation of Medicines*, a thorough and holistic account of UK drug regulation, Emily Jackson claims that a central assumption underlying her book is the idea that “there is something

distinctive about medicines which necessitates a special regulatory regime.”⁴⁷⁰ In this section, I introduce four common justifications for this special treatment of the ‘medical.’

First, consider that the centrality of health to human well-being and flourishing is widely recognized: it is a ‘primary good’ according to Rawls that all rational and reasonable citizens, regardless of their conception of the good life, desire and require at some basic level to pursue their plans and projects.⁴⁷¹ And, as Norman Daniels argues, disease or disability may greatly reduce the range of opportunities available to individuals, suggesting the importance of healthcare in ensuring fair equality of opportunity in society.⁴⁷² Universal recognition of the centrality of health to human well-being underpins the protection and promotion of public health as one of the core aims of the modern state: state action to protect public health is premised upon truly *public* reasons, that all – regardless of the conflicting content of their comprehensive doctrines – can find mutually acceptable. This public interest in health protection underwrites the legitimacy of safety regulations for both medical products and consumer products alike; the rise of the consumer protection movement and the *General Product Safety Regulations 2005* for consumer products can be understood in part as a recognition of and attempt to address the serious risks to health and life that non-medical, consumer products – like cars and electrical home goods – can pose.⁴⁷³

On this line of reasoning, one justification for a special regulatory framework for medical products that attends to *more* than just safety is because they uniquely treat ill health: medical products, on the MHRA’s definition, serve a diagnostic or therapeutic

⁴⁷⁰ Emily Jackson, *Law and the Regulation of Medicines* (Oxford: Hart Publishing Ltd, 2012): 1.

⁴⁷¹ Rawls, *Political Liberalism*, 178-187.

⁴⁷² Norman Daniels, “Normal Functioning and the Treatment-Enhancement Distinction,” *Cambridge Quarterly of Healthcare Ethics* 9, no. 3 (2000).

⁴⁷³ *The General Product Safety Regulations 2005*, no. 1803, <http://www.legislation.gov.uk/uksi/2005/1803/contents/made>.

purpose limited to disease, disability, or symptoms of pathological significance, while others do not. As Jackson describes it, “The point of the [medical] regulatory system is not just to protect the public against actively *dangerous* products,” as the regulation of all consumer products intends to do, “but also from *useless* ones” (emphases added).⁴⁷⁴ Given the central importance of health, and the recognition that medicines and medical devices are such a crucial means to protecting and improving it, we care deeply about identifying which medical products ‘work’ for various conditions and illnesses. This is one of the reasons why many, including the House of Commons Science and Technology Committee,⁴⁷⁵ have argued that homeopathic remedies – currently exempt from proving effectiveness despite making medical claims – should be subject to the same clinical standards of effectiveness as other medicinal products. Even if a product is *safe*, an individual may be exposed to harm if he fails to seek out and receive existing, *effective* treatments for his condition.

This suggests two further reasons for the special management of medical products. First, we want not only to be able to tell which medical products work but also to actively promote the development of new and better ones. Over-burdensome regulation, by inhibiting medical innovation, may prevent the development of new medical products that treat previously untreatable conditions or improve upon existing treatments. While over-regulation can undesirably stifle innovation in many different sectors, the consequences in the medical sector are considered particularly serious because of the *kind* of product innovation that it stifles, namely that which may protect or promote health. Public health, in other words, is threatened by both dangerous drugs and devices, on the one hand, and over-demanding legislation that prevents innovation of

⁴⁷⁴ Jackson, *Law and the Regulation of Medicines*, 2.

⁴⁷⁵ Science and Technology Select Committee, *Evidence Check 2: Homeopathy*, 8 February 2010, HC 45 2009-2010.

life-improving and life-saving treatment, on the other. Confirming the dual attention this requires, the European Commission's directive on medicines regulation posits, "The essential aim of any rules governing the production, distribution and use of medicinal products must be to safeguard public health. However, this objective must be attained by means which will not hinder the development of the pharmaceutical industry or trade in medicinal products."⁴⁷⁶ Medical regulation must strike a fine balance between these two public interests: protection from harmful medical products and access to innovative and beneficial ones.

The second reason for a special regulatory framework implicated by the importance of what works in the medical field is the expertise required to answer this question. As one of the WHO's main resources for designing regulations for the pharmaceutical sector observes:

Because a consumer cannot independently assess the safety, efficacy, or quality of pharmaceuticals, these products are universally recognized as being different from ordinary items of commerce, such as clothing or household appliances, and therefore in need of handling by specially trained health professionals.⁴⁷⁷

There seem to be two distinct claims about expertise in this justification for the distinctiveness of the medical sphere. First is the fairly familiar claim that consumers do not have the necessary knowledge and skills, let alone resources, to 'independently assess' medical products. Undertaking traditional judgements of safety, quality, and efficacy requires, first, the considerable resources and cooperation of pharmaceutical companies that carry out clinical trials and, second, independent clinical and scientific

⁴⁷⁶ European Commission, "Directive 2001/83/EC of the European Parliament and of the Council: On the Community Code Relating to Medicinal Products for Human Use," November 6, 2011, http://ec.europa.eu/health/files/eudralex/vol-1/dir_2001_83_consol_2012/dir_2001_83_consol_2012_en.pdf, 3.

⁴⁷⁷ Enrique Fefer, "Chapter 6: Pharmaceutical Legislation and Regulation," in *MDS-3: Managing Access to Medicines and Health Technologies* (Arlington: Management Sciences for Health, 2012), <https://www.msh.org/sites/msh.org/files/mds3-jan2014.pdf>, 7.

experts working on behalf of the MHRA or European Commission who are able to impartially evaluate the design and results of these trials. This is a claim, then, about the superiority of experts relative to the average consumer in accessing and assessing scientific evidence.

The second but related claim in this rationale is that consumers *are* able to independently assess whether ‘ordinary items of commerce’ work. On the one hand, this could refer to the perceived ability of the average consumer to adequately assess the available commercial information, consumer reports, or customer reviews in choosing, for instance, what vacuum to buy on the basis of how well it cleans floors. The skills and knowledge required to make this judgement are not so concentrated in a class of technical experts that we think it necessary or justifiable for expert regulators to make the decision of ‘what works best’ for consumers, where we do think this asymmetry justifies expert regulation of medical products. Judgements like this about when decisions should be left to consumers or assumed by experts are of course intricately bound up with the centrality of health to well-being, which makes it more important to precisely determine whether a ventilator effectively helps a patient breathe than whether a vacuum effectively cleans the living room floor. However, and keep in mind for later in this chapter, we *do* have extensive safety regulations for electrical appliances, like vacuums, on the same grounds of 1) expertise asymmetry, and 2) the importance of health: the average consumer will not have the skills or time to determine whether a vacuum’s wiring is likely to cause an electrical fire and, considering the magnitude of harm to human health, life, and property that such risks pose, the UK has specific legislation that require the use of quality assurance systems based on harmonised

European standards, which Local Authority Trading Standards Services may randomly inspect.⁴⁷⁸

On the other hand, the WHO's claim that consumers can independently assess ordinary consumer goods might be a claim about the more *subjective* nature of defining what 'works' in non-medical products. With traditional medical devices, like corrective glasses or a prosthetic leg, the standard of effectiveness may be defined relatively objectively: glasses work to the extent that they improve vision and prosthetic legs work to the extent that they increase mobility. On the margins, we may recognize less objective factors, like whether the degree of effectiveness for a prosthetic leg depends on its ability to look 'real,' but we assume that the core of the definition of effectiveness in medical products is fairly objective. This assumed objectivity is rooted again in the universally recognized value of health: we agree despite conflicting visions of the good life that health is necessary to it, that it is in a sense objectively good. Given this, there are certain dimensions within health that we also recognize as objectively good, all things considered: sight, hearing, mobility, and mental capacity, for instance. The extent to which glasses improve vision rests its claim to being an objective standard of effectiveness on the grounds that we all recognize the value of sight. The perceived objectivity of 'what works' in medical products thus assumes reasonable agreement on the value of health-related outcomes. On this account, experts can easily identify when medical products work because they share the same health-related values that we, the public and patients, do.

In contrast, the WHO's claim suggests that for many consumer goods, like cosmetics, the standard of 'what works' is recognized to be quite subjective: what it

⁴⁷⁸ See, for example, Business Companion, "Electrical Equipment: Essential Information for Suppliers of Electrical Equipment, Including CE Marking," In-Depth Guide, <http://www.businesscompanion.info/en/quick-guides/product-safety/electrical-equipment#Satisfyingthegeneralsafetyrequirement>.

means for a hair gel or make up to be effective is interpreted far more broadly than for medical products and, furthermore, the *value* placed on these already subjective benefits is likely to vary significantly between different individuals. You may think that mascara is effective when it makes eyelashes longer, I may think it is effective when it makes them darker, and our friend may agree with you over me but place very little value on this ‘lengthening benefit’: we might end up buying two different types and our friend might buy none at all, but none of us is necessarily wrong and certainly not unreasonable in this disagreement, as we think the standard of effectiveness here is subjective. In such cases, where the benefits of a product are largely subjective, we think it desirable to allow as much consumer choice as possible within some bounds of safety, because the individual consumer is much better placed than experts to determine whether such products are ‘effective.’ In contrast, we think medical products have largely objective measures of effectiveness and on that basis we are not as concerned about experts – rather than patients themselves – making decisions about effectiveness for us. Because they are merely calculating on the basis of objective standards of health, we would decide the same way the experts decide if we had the information, skills, and time the experts had, the argument goes. The argument from expertise for a distinct medical regulatory pathway thus rests on a complex pair of premises. All things considered, expert intervention in citizens’ decision-making is more desirable the more technically complex and resource intensive is the evaluation of the product’s effectiveness. But, equally, expert-determined standards of efficacy for the population will be less desirable the more subjective is the criteria for reasonably defining effectiveness.

So far, I have pointed to four reasons, albeit intimately connected, given for why medical products are different from and thus regulated on separate grounds from non-medical products – why, in Jackson’s terms, there is something *distinctive* about the

medical: 1) given the value of health, the need to identify which medical products work, 2) given the value of health, the need to promote innovation of new medical products that work, 3) the relative amount of resources and degree of scientific expertise needed to determine when medical products are safe and effective, and 4) the relative objectivity of what it means to be effective for medical products. In the following section, I hope to show through consideration of the case of transcranial direct current stimulation (tDCS) devices, which have both therapeutic and non-therapeutic purposes, how these four justifications for the distinctiveness of the medical from the non-medical do not always hold up.

3. Cognitive enhancement: benefit beyond ‘medical’ benefit

The borderline problems posed by transcranial direct current stimulation (tDCS) challenge some of the justifications for treating medical and consumer products differently. TDCS devices function by sending low intensity currents between electrodes placed on the scalp to modulate neuronal activity. Studies have suggested potential, modest benefits of tDCS for both therapeutic purposes, such as the treatment of depression, dementia, chronic pain, stroke, and Parkinson’s,⁴⁷⁹ and non-therapeutic, enhancement purposes, such as the improvement of cognitive abilities like attention, working memory, and mathematics.⁴⁸⁰ On the current definition of medical device, a tDCS device intended for treatment would receive substantially greater regulatory scrutiny than one intended for enhancement: a manufacturer producing a tDCS device

⁴⁷⁹ Lilly Knechtel, Renate Thienel, and Ulrich Schall, “Transcranial Direct Current Stimulation: Neurophysiology and Clinical Applications,” *Neuropsychiatry* 3, no. 1 (2013); Agnes Flöel, “tDCS-Enhanced Motor and Cognitive Function in Neurological Diseases,” *Neuro-Enhancement* 85, part 3 (2014); F. Fregni et al., “Regulatory Considerations for the Clinical and Research Use of Transcranial Direct Current Stimulation (tDCS): Review and Recommendations from an Expert Panel,” *Clinical Research and Regulatory Affairs* 32, no. 1 (2014).

⁴⁸⁰ For a broad overview, see Roi Cohen Kadosh, “Using Transcranial Electrical Stimulation to Enhance Cognitive Functions in the Typical and Atypical Brain,” *Translational Neuroscience* 4, no. 1 (2013).

for a therapeutic purpose will have to perform a clinical assessment prior to marketing it with substantially greater oversight both before and after it comes on the market, while a manufacturer of a tDCS device for a cognitive enhancement purpose will only have to meet general product safety requirements that are only enforced *ad hoc* after the product has gone on the market. However, tDCS for treatment and tDCS for enhancement pose similar *risks* and function to produce cognitive *benefits* in a similar way.⁴⁸¹

Let us first think about the risks that tDCS devices pose. The concept of risk depends upon both the magnitude of potential harm and the probability of experiencing it.⁴⁸² When used with recommended current levels and protocol, a group of international experts summed up the known adverse effects of tDCS as “mild, benign and short-lived,” with the most common adverse effects being short-term skin reddening and itching.⁴⁸³ Indeed, they found no serious adverse events – such as cardiorespiratory arrest, seizures, or brain damage – in over 10,000 subjects examined in the tDCS literature between 1998 and 2014. However, the expert panel is careful to note that the minimal risk posed by tDCS devices depends upon it being used under clinical supervision and within ‘standard parameters’ – including a maximum current level, the use of electrodes that minimize skin burns, an application period of less than 20-60 minutes, and no more frequent use than twice per day.⁴⁸⁴ As such, build-it-yourself as well as commercial, pre-assembled devices cannot be assumed to have the same safety profile.

Risk based on what we might call the *inherent* properties of a tDCS device will be the same regardless of whether it is used for a therapeutic or enhancement purpose.

⁴⁸¹ Maslen et al., “The Regulation of Cognitive Enhancement Devices.”

⁴⁸² Charles Weijer, “The Ethical Analysis of Risk,” *The Journal of Law, Medicine & Ethics* 28, no. 4 (2000).

⁴⁸³ Fregni et al., “Regulatory Considerations for the Clinical and Research Use of Transcranial Direct Current Stimulation (tDCS),” 2.

⁴⁸⁴ *Ibid.*

These risks inherent to the device (rather than dependent on how it is used) include electrical malfunction resulting in delivery of current levels that are too high or the use of suboptimal material for the electrodes. In contrast, risks stemming from the device's *interactional* properties, such as the context in which it is used and the person administering it, are likely to vary between therapeutic and enhancement use. Devices used for therapy will tend to be administered directly by a clinician or at least following clinician instruction, which should reduce the risk of using the device for too long or too frequently. Devices for enhancement, either bought online or constructed at home, will be used by consumers who may not be aware of the standard parameters for tDCS that appropriately minimize risk. Because they are not 'medical,' however, neither the inherent risks from faulty electrical design and poor materials, nor the interactional risks from uninformed use, would be subject to MHRA oversight on its traditional understanding of its regulatory bounds. Ironically, then, the use of tDCS devices for enhancement may pose *greater* traditional health risks to individuals than those used in therapy. As such, the importance of protecting health cannot justify subjecting the 'consumer' tDCS product to such significantly reduced risk assessment.

The MHRA seems to be recognizing how this distinction between devices that serve a medical purpose and those that do not does not always adequately suit its responsibility of protecting public health. The ongoing revision of the Medical Device Directive proposes to widen the scope of medical device regulation to include "certain implantable or other invasive products *without a medical purpose* that are similar to medical devices in terms of characteristics and risk profile (e.g. non-corrective lenses, implants for aesthetic purposes)."⁴⁸⁵ Because they are not technically 'implantable' or

⁴⁸⁵ European Commission, "Proposal for a Regulation of the European Parliament and of the Council: On Medical Devices, and Amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation

‘invasive,’ it is not assured that this revision will cover tDCS; however, it seems quite likely considering that tDCS devices for enhancement are indeed products without a medical purpose with similar characteristics and risk profile as tDCS devices for therapy (the exact same device could be used for both purposes). This adjustment to the MHRA’s regulatory framework, the relaxing of its sharp distinction between the medical and the consumer, would appropriately recognize that certain consumer products have similar or the even the same characteristics and risks as medical products. However, it also raises the question of how ‘benefits’ and ‘effectiveness’ should be conceptualized for such non-therapeutic devices.

It is notable that tDCS for enhancement displays, in addition to similar health risks, characteristics of two of the other main justifications for regulating medical products separate from ‘ordinary consumer’ ones. Although – as both Maslen et al. and Andreas Kuersten and Roy Hamilton note – the benefits of cognitive enhancement will be *valued* differently by different individuals,⁴⁸⁶ there is still a relatively objective concept of when cognitive enhancement devices ‘work.’ The benefits and effectiveness of tDCS as therapy are traced to, for instance, improvement in cognitive function in Alzheimer’s patients and motor or language function recovery in stroke patients, fairly objective standards of effectiveness that we can compare to other treatments that seek similar outcomes. The benefits and measure of effectiveness for tDCS for enhancement are similarly based on cognitive function, just in ‘healthy’ individuals, caused by similar neurological effect; the benefits are not so subjective, as with make up or clothing, that we cannot conceptualize effectiveness. Second, and more importantly, the average

(EC) No 1223/2009” (Brussels, September 9, 2012), http://ec.europa.eu/growth/sectors/medical-devices_old/documents/revision/files/revision_docs/proposal_2012_542_en.pdf, 4.

⁴⁸⁶ Andreas Kuersten and Roy H. Hamilton, “The Brain, Cognitive Enhancement Devices, and European Regulation,” *Journal of Law and the Biosciences*, (2014); Maslen et al., “The Regulation of Cognitive Enhancement Devices.”

consumer cannot independently assess the safety and effectiveness of tDCS, regardless of whether it is for medical or consumer use: this requires access to substantial clinical trial data and manufacturer design assessments, as well as clinical and scientific expertise on this information. In other words, in order to benefit from the use of tDCS for enhancement, consumers need access to the expertise and data that the MHRA *already holds* in relation to tDCS for treatment. The MHRA is the regulator most well-placed to assess the risks and benefits of these products, even when they are ‘non-medical.’

However, returning to Maslen and colleagues’ argument, although experts are better placed “to determine the nature, size and probability of the effects of [tDCS]...consumers are best placed to evaluate the impact of these (expert-identified) effects on their own wellbeing.”⁴⁸⁷ This argument captures the fact that the weighing of benefits against the risks of a product, and thus the informed decision to use it, will vary between individuals based on the degree to which they value those benefits, like improved working memory. It suggests the important role of accurate and understandable user information leaflets, similar to those provided for traditional medicines, that would impart to consumers expert-verified accounts of the risks and benefits but leave the decision about whether the risk-benefit trade-off is acceptable to the user. This does not mean that dangerous devices should be approved so long as all the information is provided, as Maslen et al. point out, but rather that “in cases where it is contested whether the benefits make the risks acceptable – where there is room for reasonable disagreement – the device should be approved so that the consumer can adjudicate for herself.”⁴⁸⁸ In other words, a broader *range* of risk-benefit ratios should be

⁴⁸⁷ Maslen et al., “The Regulation of Cognitive Enhancement Devices,” 88.

⁴⁸⁸ *Ibid.*

accepted as reasonable in licensing decisions about tDCS for enhancement, given that people will value the benefits differently and should be allowed to, and supported in, making such autonomous decisions on different sets of values or preferences given a range of conceptions of the good life.

This consideration of tDCS devices has shown how three of the four justifications for medical exceptionalism will not always hold: products that do not aim to treat disease may raise similar risks by similar mechanisms, entail relatively objective rather than subjective measures of effectiveness, and require the assistance of the same types of scientific expertise and data to reach an informed decision about risks and benefits of use. Of course, measuring effectiveness for the purpose of licensing tDCS for specific therapeutic uses will still be a distinct process: a tDCS device could have a favourable risk-benefit profile – because the risks are minimal and it provides a demonstrable benefit by improving working memory, for example – without being clinically effective – because this benefit has little or no bearing on the condition it is being considered for and the outcome measures we identify, say stroke victims for whom we are seeking effective speech therapy.

Furthermore, this case has illustrated how consumer products may provide ‘non-medical’ or ‘non-health’ benefits that are nevertheless measureable and require scientific expertise and access to trial results to assess. Because users will place different values on these non-medical enhancement benefits, the MHRA must expand the consideration of ‘acceptable’ risk-benefit profiles to identify a range based on different but reasonable valuations. Simply increasing the MHRA’s expertise will not suffice to make these judgements about what valuations are reasonable, because to be publicly accountable such decisions must be made on the basis of reasons that those who will have to live by them can accept. The MHRA must thus reform its licensing process, particularly its

concept of public and patient/user involvement, to set these ranges publicly. I will expand on how this might be accomplished in the final section of this chapter.

Although it may be particularly clear in the case of tDCS for enhancement how different people might value the benefits differently, this feature of MHRA's judgements is not limited to enhancement or even to novel biotechnologies. Rather, the MHRA does and has always had to make decisions about when risk-benefit trade-offs are 'acceptable'; value judgements are thus ubiquitous in the MHRA, not unique to these borderline cases. As discussed in the previous chapter, it is likely that many health-related benefits of treatments – to the quality and length of life – are recognized as highly valuable by all, making reasonable estimations of their value by experts easier and less contentious: a reduction in pain or the extension of life are things that all of us recognize, all else equal, as beneficial. But particularly given treatments in which these valued benefits are inversely correlated, for example, a cancer treatment that extends life but also increases pain or reduces mobility, patients will have conflicting but reasonable positions on whether the risk-benefit tradeoff is acceptable given different values placed on expected health gains and health losses in different dimensions of quality of life. Recall that for NICE's resource allocation decisions, where allocation of limited public funds is at stake, the value placed on changes in different dimensions of quality of life is based on averages from a representative sample of the public (the EQ-5D method). In the MHRA context, the question is not whether the NHS should fund a drug but whether it should be available, full stop, in the UK for public or private purchase. Because of this, a reasonable *range* of how actual users may consider the risks and benefits must be the basis for judgements of what risk-benefit profiles are 'acceptable,' rather than an average from the public or an estimation from experts who may, for instance, over-weigh the value of traditional health gains.

Recognizing that these judgements about the value of different medical, or non-medical, benefits may determine whether or not citizens will be coercively prevented from accessing a drug or device, the process of setting ‘acceptable’ risk-benefit ratios and definitions of effectiveness must be based upon reciprocal reasoning in order to be publicly accountable. This puts patient and user values at the centre of decisions, which in turn, as the final section will lay out, calls for procedural changes to patient and public involvement. Before these recommendations, however, I examine the borderline problems posed by telehealth systems, which are useful in illustrating a second but related way in which the medical cannot be so easily roped off from the non-medical, and what this entails for the MHRA.

4. Telehealth systems: risk beyond ‘health’ risk

The MHRA borderlines between different ‘classes’ of medical devices are based on the level of risk that devices pose to patients. As long as a product meets the two definitional prongs of intended purpose and function set out in the first section of this chapter, it will fall into the category of ‘medical device’ regardless of whether it is for use by patients in the home or for administration by a doctor in the clinical context. Because of the wide net that this definition casts, the category of medical devices encompasses an incredibly diverse range of products intended for use by a variety of actors, from plasters, thermometers, and syringes to x-ray machines, pacemakers and HIV blood tests. As such, while the definition provides an explanation of these products’ common *nature*, further demarcation within the category of devices attempts to group this varied set of products using a risk-based system.

Within this risk-based system, devices are classified into Classes I, IIa, IIb, and III, with Class I encompassing the lowest risk devices and Class III containing the

highest risk devices. Each class has a distinct licensing pathway: while Class I devices may be self-certified by the manufacturer, Class III devices must be certified by a Notified Body and include an assessment of the design and manufacturing stages due to their high risk. Classification is thus crucial in the MHRA's regulatory process, as the higher the class of a device, the higher the standards that the device's manufacturer must meet in order to prove that the risks are acceptable. The European Commission suggests that it follows "a 'risk based' system *based on the vulnerability of the human body* taking account of the potential risks associated with the devices" (emphasis added).⁴⁸⁹ If we turn to the Commission's guidance for classification, there are 18 complex 'categorization rules' for determining which class a device falls into. In particular, the Commission's guide to classification speaks of invasiveness, time, and organic function as the three "normal criteria" of human vulnerability that impact the risks associated with a device. Generally speaking, in other words, a device will entail higher risks given the nature of the human body, 1) the more invasive it is, whether invasive in a bodily orifice or surgically invasive, 2) the longer it will remain in contact with, implanted within, or acting upon the human body, and 3) to the extent that it acts upon vital systems, like the nervous or cardiovascular systems, particularly by incorporating a medicinal product.

If we turn to the discussion of risk within the Medical Devices Directive's Essential Requirements that all devices must meet, we are told that at the most general level:

[D]evices must be designed and manufactured in such a way that, when used under the conditions and for the purposes intended, they will not compromise the clinical condition or the safety of patients...provided that any risks which may be associated with their intended use constitute

⁴⁸⁹ European Commission, "Medical Devices: Guidance Document - Classification of Medical Devices," Guidance relating to the application of the Council Directive 93/42/EEC on medical devices (Brussels, June 2010), http://ec.europa.eu/health/medical-devices/files/meddev/2_4_1_rev_9_classification_en.pdf, 4.

acceptable risks when weighed against the benefits to the patient and are compatible with a high level of protection of health and safety.⁴⁹⁰

Risks must be eliminated or reduced where possible, and, if not possible, protection measures must be put in place and users must be informed of “residual risks,” with any “undesirable side-effects” included in the risk-benefit weighing.⁴⁹¹ Crucially, these Essential Requirements highlight that the assessment of acceptable risk requires attending to “the environment in which the device is intended to be used” and “the technical knowledge, experience, education and training and where applicable the medical and physical conditions of intended users.”⁴⁹² In other words, both the context in which a device is used and the characteristics of the person using it will affect the risks it presents.

On the one hand, then, the classification rules determining which assessment standards a device must meet suggest that it is mainly the *inherent properties* of a device that make it more risky: devices whose intended use is inherently invasive, long term, or directed at a vital organ like the heart will be classed higher than those that are not and thereby receive greater scrutiny of risks and benefits. On the other hand, the Essential Requirements suggest the importance of risks that stem from the *interactional properties* of devices: holding inherent properties constant, the riskiness of a device is affected by the context in which it is used – the GP’s office vs. the intensive care unit vs. the home – and the person responsible for carrying out its intended use. There seems to be a mismatch, then, between the classification system’s focus on inherent properties and the recognition that risks are affected by interactional properties as well. Traditionally, this

⁴⁹⁰ European Commission, “Council Directive 93/42/EEC Concerning Medical Devices,” June 14, 1993, <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:1993L0042:20071011:EN:PDF%20,25>.

⁴⁹¹ Ibid.

⁴⁹² Ibid.

has not been a problem, as devices with high-risk inherent properties are precisely those whose risks are mitigated by the context in which they are used and the persons responsible for administering them: invasive devices like cardiovascular catheters or joint replacements are administered in the highly ordered and monitored hospital environment by a trained surgeon. Devices traditionally used in the home and which may be bought by average consumers in stores or over the counter, on the other hand, are low risk, like plasters, thermometers, and glasses.

This is the point at which the increasing use of more complex and life-saving medical devices in the home, as opposed to the clinic, sheds light on the shortcomings of the MHRA's traditional processes for assessing risks. As suggested in Chapter II, this trend toward the use of medical devices in the home is most common in elderly patient who rely upon a diverse set of telehealth devices or systems. Consider first how elderly patients may now measure, record, and transmit their vital signs to a doctor or remote caregiver, using active diagnostic devices previously reserved for the clinic that are connected to phone or computer apps. Clearly the inherent properties of, for instance, an electronic blood pressure monitor are not altered by its relocation to the home: it is used on the same part of the body, for the same amount of time and is no more invasive than when it is used in a GP's office. But the risks – and indeed benefits – are altered by the changes this move induces in their interactional properties. Patients may find it more difficult to read or understand their results, although this may be a fairly straightforward issue of re-designing a device for patient rather than professional use. More fundamental, however, is the change in context: patients alone in the home, particularly those with no family or regular carers, rely upon the effectiveness and accuracy of such diagnostic devices to reap both traditional 'medical' benefits of health protection and non-medical benefits to their well-being, like increased independence. This kind of telehealth system

effectively replaces what would have once been far more regular contact with a health professional. As such, were the measuring or reporting functions or threshold alarms of the device to fail, the patient may remain unaware of his need for medical care, need that may previously have been caught by a doctor.

Currently, “active devices for diagnosis” like electronic blood pressure monitors are considered lower-risk, Class IIa devices, while “active devices for diagnosis *where variations could result in immediate danger to patient*” are placed in higher risk Class IIb, like intensive care monitoring devices.⁴⁹³ Yet because telehealth vital sign monitoring devices do not present *immediate* danger, they are considered lower risk:

“Medical devices intended to be used for continuous surveillance of vital physiological processes in anaesthesia, intensive care or emergency care are in Class IIb, whilst medical devices intended to be used to obtain readings of vital physiological signals in routine check ups and in self-monitoring are in Class IIa.”⁴⁹⁴

The classification system thus seems to overlook the evaluation of a different set of risks associated with the move to home care and reliance on at-home medical devices, including the decreased oversight by medical professionals and the vulnerability of elderly patients who may be at home alone when acute episodes that might have been predicted hit.

This is a very different sort of vulnerability than the ‘vulnerability of the human body’ focused upon by the MHRA in device classification and risk assessment: it stems not necessarily from any weakness of the human body – although lack of mobility and chronic conditions in the elderly certainly fuel the problem – but rather from the home context. Recognizing the risks from this distinct kind of vulnerability requires attention

⁴⁹³ European Commission, “Medical Devices: Guidance Document - Classification of Medical Devices,”

42.

⁴⁹⁴ Ibid., 43.

to a broader range of non-medical considerations than the MHRA is accustomed to, ones that follow from the *social* context of the home.

Special attention by the MHRA to such extra-ordinary considerations is not, however, unprecedented. In describing why it sets out a special high-risk classification rule for devices used for contraception or prevention of sexually transmitted disease, the European Commission notes that these two “intended uses relate to special cases of human vulnerability that cannot be covered by the normal criteria of time, invasiveness and organic function.”⁴⁹⁵ While some traditional health risks associated with STDs and pregnancy play into this ‘special case’ of human vulnerability, it also seems based upon the recognition of the special *social* implications of sex, which introduces risks that are more than just ‘medical’ in nature. The risks from the failure of a contraceptive device, for instance, cannot be wholly captured by health considerations: an unintended pregnancy may negatively impact well-being for social and personal reasons as much as the health risks associated with an abortion or birth.

Telehealth systems, based in the social context of the home, may also be considered a special case where the true extent of vulnerability and risk are not captured by the narrow focus on medical risks traced to the bodily impact of invasiveness, time, and area of focus in question. The cases of GPS tracking devices for Alzheimer’s patients and fall detectors for elderly patients living alone, two other increasingly popular telehealth devices, are particularly apposite in this respect. The MHRA notes that “apps with software that monitors a patient and collects information entered by the user, measured automatically by the app or collected by a point of care device may qualify as a medical device *if the output affects the treatment of an individual* (emphasis

⁴⁹⁵ Ibid., 47.

added).⁴⁹⁶ Both devices seem to meet the definition of a medical device as their outputs affect the treatment of patients, although neither is currently consistently regulated as such. A GPS device for an Alzheimer's patient may be a crucial part of a treatment plan for a patient who requires frequent medication, and a fall detector may ensure that a patient receives appropriate emergency care if injured and unable to call for help.

Both devices present key risks and benefits based on the context of their use and their user that are not strictly health-related. Unlike the example of home measurement and telereporting of vital signs, neither is a case of a device previously used in the clinic moving to the home. Yet each does fulfil a role of monitoring that previously would have required constant observation and containment precautions by doctors or carers, often in medical institutions. Indeed, similar to home monitoring of vital signs, the benefits of GPS trackers and fall detectors may largely be traced not to any technical measurements of improved health but to increased well-being from greater independence and control over one's life or the ability to remain at home with loved ones rather than in a care facility. However, these benefits are accompanied by risks that are similarly non-medical in nature. Much attention has been dedicated to the ethical issues surrounding the use of GPS tracking devices in patients with dementia.⁴⁹⁷ For patients with more mild forms of dementia who are still capable of giving consent in this type of decision, easy purchase and home administration of such devices raises the risk that patients' autonomy and privacy will be violated if their refusal to use them is not respected by family or carers. And, for patients lacking capacity to consent in this case, the isolated home

⁴⁹⁶ Medicines and Healthcare Products Regulatory Agency, "Medical Device Stand-Alone Software Including Apps."

⁴⁹⁷ See, for example, L. Robinson et al., "Balancing Rights and Risks: Conflicting Perspectives in the Management of Wandering in Dementia," *Health, Risk & Society* 9, no. 4 (2007); J. C Hughes, "Electronic Tagging of People with Dementia Who Wander," *BMJ* 325, no. 7369 (2002); S. Welsh et al., "Big Brother Is Watching You--the Ethical Implications of Electronic Surveillance Measures in the Elderly with Dementia and in Adults with Learning Difficulties," *Aging & Mental Health* 7, no. 5 (2003).

context means that there will be less oversight in judgements about the patient's 'best interests,' which, as elder care specialists Julian Hughes and Stephen Louw stress in raising such ethical concerns, ought not be limited to *medical* best interests.⁴⁹⁸

With respect to fall detectors, a set of devices which include both automatic sensors and video monitoring equipment that are used to report falls to either health and care professionals or family and friends, Michelle Ganyo, Michael Dunn, and Tony Hope have introduced the concept of "ethical side-effects" that companies should be required to make explicit to users.⁴⁹⁹ While some of these ethical side effects implicate traditional health risks, like the chance that the device will *increase* rather than decrease the risk of falls if patients switch them off without carers' knowledge, a bulk relate to privacy and autonomy, as in the GPS tracking device case. As an especially difficult case of risks that extend beyond health, the authors give the example of a woman with mild dementia who agrees to video monitoring that will be sent to select family members: "Because of the dementia she readily forgets that some of her family can see exactly what she is doing. In the privacy of her bedroom she behaves in ways that she would not want her family to observe. Was her consent to have the system installed valid?"⁵⁰⁰ More maliciously, family members and carers could use such devices to socially control and contain patients for personal gain, like a relative attempting to extort money, for instance.

The MHRA attempts to distinguish between different uses of the same kind of device, saying:

⁴⁹⁸ Julian C Hughes and Stephen J Louw, "Electronic Tagging of People with Dementia Who Wander," *BMJ* 325, no. 7369 (2002).

⁴⁹⁹ Michelle Ganyo, Michael Dunn, and Tony Hope, "Ethical Issues in the Use of Fall Detectors," *Ageing & Society* 31, no. 08 (2011).

⁵⁰⁰ *Ibid.*, 1356-1357.

[C]onsideration should be given to the difference between social care, well-being and health, which can become blurred. For example, an app that uses an accelerometer or gyroscope as a falls detector in epileptic patients is likely to be regulated as a medical device but the same app or device could alert someone if an elderly person has got up from a chair or bed in a social care context. As a detector of falls of a medical condition the app will qualify under the MDD and be regulated as a medical device but in the second case it will not meet the definition of a medical device and the medical device regulation would not apply.⁵⁰¹

It is an odd example as it seems clear that a fall detector for use by an elderly patient with, for instance, osteoporosis or knee and hip injuries, fits just as much in the definition of medical device as a fall detector for epileptic patients: both are intended for the monitoring, treatment, and alleviation of a disease, injury or handicap. But the more troubling implied distinction is between a fall detector used to socially surveil a patient's movements (getting up from a chair or bed) and a fall detector used for medical monitoring (reporting falls to ensure help arrives). These are indeed distinct uses, and importantly so, of telehealth devices, but it is the same device in question that may be utilized for either or both purposes. They cannot be concretely decoupled when such a device is used in the uniquely isolated and relatively unregulated social space of the home. Because the non-medical risks to well-being, be they to privacy or independence or dignity, are inextricable from the device regardless of its intended purpose, the potential for these risks needs to be considered in the regulation of both purposes of the device. Identification, minimization, and proper indication of these non-health risks from home monitoring and telereporting are crucial in safeguarding elderly patient's well-being. Attention to such 'ethical side-effects' should thus be fed into the Essential Requirements' more general mandate that "Any undesirable side-effects must constitute

⁵⁰¹ Medicines and Healthcare Products Regulatory Agency, "Medical Device Stand-Alone Software Including Apps."

an acceptable risk when weighed against the performances intended,” even though this has traditionally been reserved for health-related side effects and risk.⁵⁰²

In sum, the home use of telehealth systems highlights how medical devices often implicate non-medical considerations of risk and benefit. Because of the significant way in which these considerations may impact patient well-being, the MHRA’s attempts to delimit the ‘medical’ aspects in risk assessment and classification of medical devices is not fit for purpose. Health, as discussed in section one, is a primary good universally recognized by reasonable persons despite conflicting accounts of what it means for one’s life to fare well. But this is largely, if not exclusively,⁵⁰³ because good health enables and enhances well-being while ill health may seriously hinder it. As such, when other aspects of a device – apart from its medical risks and benefits – may seriously impact well-being, it will not do to separate them out on grounds of the distinctive importance of health to well-being or flourishing. There is a general recognition in the Medical Devices Directive of the importance of the context in which a device is used and the characteristics of the person using it in assessing its risks. But because the risk-based classification system depends so heavily upon inherent properties of devices that are closely tied to bodily harm or health risks, the classification of devices that present alternative but serious kinds of non-health risks based largely on interactional properties may miss these risks. Furthermore, important non-medical benefits to well-being that weigh in favour of approval – such as increased independence – may be overlooked by a narrow focus on a product’s ability to deliver medical benefit.

⁵⁰² European Commission, “Council Directive 93/42/EEC Concerning Medical Devices,” 25.

⁵⁰³ On the debate over whether the value of health is intrinsic or instrumental, see Peter Duncan, “Health, Health Care and the Problem of Intrinsic Value,” *Journal of Evaluation in Clinical Practice* 16, no. 2 (2010).

5. Implications for the processes and regulatory boundaries of the MHRA

The bulk of this chapter has been dedicated to breaking down the conception of the MHRA as a purely technical, standard-meeting arm's length body. Despite attempts to cordon off the area of the 'medical' from the 'non-medical,' 'social,' or 'consumer,' consideration of two novel sets of technologies – tDCS and telehealth systems – exemplify both the fluidity of these boundaries and the potential for reasonable disagreement about the valuation of benefits against risks (both those that are traditionally recognised as 'medical' and those that are not). This suggests two key conclusions. First, casting the MHRA's licensing process as one based solely on 'technical' analysis belies important value judgements about risks and benefits that end up affecting the treatments that citizens may access. Second, on account of the fluidity insight, devices or drugs may implicate non-medical risks and benefits to the user's wellbeing that cannot be decoupled from its medical use, and capturing these wider considerations requires attending to the holistic effect of a product. This may mean considering dimensions commonly understood as 'social' or otherwise beyond the MHRA's scope. Although particularly vivid in the area of emerging biotechnologies, both conclusions hold in the licensing of more 'ordinary' medical treatments as well.

As such, the MHRA as a case study has uniquely allowed me to demonstrate why my public accountability framework must be applied even within a context that initially appears quite distinct from that of the HFEA and NICE, the two that have received the lion's share of attention on questions of public reasoning and deliberation. It thus serves within the wider context of the dissertation as an example of the prerequisite perception needed to realize that the claims I make about public accountability apply: we must be able to 'see' the true nature of institutional characteristics and decision features in order to understand that the demands of reciprocity kick in. Short-sightedness in this respect

may mean getting the ‘solution’ to problems wrong, like the MHRA’s attempt to solve borderline problems by increasing clinical expertise. Because these problems relate to the endorsement of values at arm’s length on behalf of the public, reciprocal reasoning rather than (or at least in addition to) greater technical expertise is called for.

While this chapter has thus focused on ‘earlier’ questions in the practical application of my public accountability framework, we can also identify some general implications for reforming the processes of the MHRA. First, the MHRA’s relationship to the public and specific patient groups must be adjusted. Currently, formal pathways for patient involvement are fairly limited to monitoring mechanisms that make it easier for patients to report adverse drug reactions of already-licensed medicines, like the Yellow Card Scheme.⁵⁰⁴ Beyond this, engagement with the public tends to be narrowly conceived as developing better ‘risk communication,’ with the primary remit of the Patient and Public Engagement Expert Advisory Group listed as the “development of effective communications for patients, the public and carers to help them make informed choices about medicines and to use medicines safely.”⁵⁰⁵ However, it is the values of those who will or want to be able to use products that must form the basis of such safety judgements; a focus solely on communicating (expert-determined) risks and benefits is a ‘one-way’ method of communication that does not recognize or allow the need to involve patients in these earlier stages.

⁵⁰⁴ Medicines and Healthcare Products Regulatory Agency, “The Yellow Card Scheme: Guidance for Healthcare Professionals,” January 23, 2015, <https://www.gov.uk/the-yellow-card-scheme-guidance-for-healthcare-professionals>.

⁵⁰⁵ Medicines and Healthcare Products Regulatory Agency, “Patient and Public Engagement EAG,” *Membership - Commission on Human Medicines*, accessed August 19, 2015, <https://www.gov.uk/government/organisations/commission-on-human-medicines/about/membership#patient-and-public-engagement-eag>.

A June 2015 conference between the MHRA and the UK BioIndustry Association suggests a growing acknowledgment of this idea; one of the recommendations coming out was that:

Risk : benefit should be seen from the patient perspective; regulators must ensure they understand what level of risk patients are prepared to accept. These assessments must factor in not only the extent to which patients' views of risk : benefit are different from those of regulators, but also that patients' views differ from one group to another, and may also vary for one individual over time, depending on the severity of the disease.⁵⁰⁶

However, processes have not yet been proposed that would allow this patient perspective to drive reasoning about safety and effectiveness. Similar to argument in the NICE chapter, patients for whom drugs or devices are intended must be involved early in the licensing process to ensure that the 'endpoints' or factors being measured in clinical trials actually capture what this set of patients care about. This means consulting with them both in the approval of trial designs and again when those trials results are returned to ensure the risk-benefit calculus reflects patients' perspectives rather than regulators' or manufacturers' perspectives.

Although there are now 'lay members' and 'patient representatives' sitting on the Commission on Human Medicines and its Expert Advisory Groups, which advise the MHRA on medicines licensing, such membership diversification is of limited usefulness for reasons similar to patient experts in NICE; a cancer patient will not be able to 'represent' either the values of, for example, an HIV patient, or even those of all cancer patients. Involving a patient or carer with experience of cancer on the Oncology and Haematology Expert Advisory Group and one with experience of HIV or AIDS on the Anti-infectives, HIV and Hepatology Expert Advisory Group would go some way to

⁵⁰⁶ UK BioIndustry Association and Medicines and Healthcare Products Regulatory Agency, "Pathway of Innovation from Research to Patients Conference" (London, June 11, 2015), http://www.biamhraconference.org/Conference_report_04_web.pdf.

keeping the right set of experiences and values in mind. But to properly elicit the range of reasonable interpretations and trade-offs of risks and benefits, the MHRA should formally consult with patient interest groups at these two key stages (trial design and assessment of the acceptability of risk-benefit profiles), who are better placed to reach out to and collect views from the spectrum of patients. This might take a similar form to the ‘written evidence submissions’ in NICE technology appraisals, and patient groups might likewise be encouraged to collect evidence from patients in a representative and methodologically robust manner.

As described earlier, it is reasonable to think that the drugs licensed to be on offer should respond to a greater range of patient characteristics and patient values than those considered by NICE. In NICE, where resource scarcity is crucial, the value of different health states is determined by an average drawn from a representative measure of the public. But the MRHA should consider that certain outcomes will matter much more or much less to individual patients, and attempt to respond to the need for such variety in treatment options by permitting a reasonable range of risk-benefit weighings: some individuals will greatly value certain quality of life improvements even if it means shorter life expectancy, while others will be willing to accept reduced quality of life for longer life. Licensing decisions, about what products are safe and effective, determine the range of options that are available both in private healthcare and as options for the NHS. As such, there is a ripple effect from this earliest stage of licensing, through to NICE technology appraisal, and finally to the individual patient consultation room: if licensing decisions identify the wrong measures of effectiveness or the wrong range of risks and benefits, this will have a knock-on effect for both these later stages of patient-centred healthcare.

This highlights the need for the reform of the MHRA Devices Division to look beyond the enhancement of clinical expertise. Because the devices field is far more diverse and undergoing more rapid development than the medicines field, there are not yet standing Expert Advisory Groups for different types of devices and it may be difficult to classify them as has been done for the Commission on Human Medicines. Furthermore, conformity assessments for medical devices are carried out by private sector Notified Bodies, rather than the MHRA itself, which introduces another layer of complexity; creating the conditions for reasoning from the patient perspective will require bringing these Notified Bodies up to date on these institutional goals. It may be that patient involvement must be arranged on a more *ad hoc* basis in the area of devices due to these features. So, in light of the concerns raised about the risk of device use in the home context, the MHRA could form a temporary working group to investigate the full range of medical and extra-medical risks and benefits in telehealth systems, whose findings could then be fed back to the Notified Bodies who are responsible for individual assessments.

This idea relates to a second key reform that follows from my claims in this chapter, one having to do not just with patient involvement processes within the MHRA but with its institutional boundaries. Recognizing the need to look beyond the traditional medical implications of drugs and devices means that the MHRA must cultivate stronger institutional relations with the bodies with whom it shares regulatory ‘borders.’ Identifying and assessing aspects of the home care setting, for instance, is an expertise of the Care Quality Commission (CQC), which regulates the provision of health and social care. It has experience going into homes and understanding the realities of this context and the use of devices within it. Partnering with the CQC and, for example, involving it in a working group on telehealth systems, would help to flag up the full range of relevant

considerations for the licensing of these devices. Similarly, the national Competition and Markets Authority as well as Local Authority Trading Standards Services, both operating under the Department for Business, Innovation and Skills, are responsible for consumer protection.⁵⁰⁷ They have the infrastructure in place to identify when devices being sold – which may not have received any pre-marketing assessment – raise safety or effectiveness concerns that may require expertise or data possessed by the MHRA to properly address (as in the case of tDCS). Partnership with these bodies can help both protect public health and promote consumer choice by ensuring that the relevant range of considerations is identified by bodies that are likely to ‘see’ products and the concept of ‘acceptable risk’ in a very different light from the MHRA.

Conveniently, there already exists a Consumer Protection Partnership run by the Competitions and Markets Authority to “ensure that partners work together effectively and important issues are tackled and do not fall between partners in the consumer landscape due to differing accountabilities,” which includes bodies like the Local Authority Trading Standards Services, Trading Standards Institute, and Citizens Advice.⁵⁰⁸ Similarly, there exists a National Collaboration for Integrated Care and Support, bringing together bodies like the CQC, NICE, and charities dedicated to representing patient voices in health and social care: “We need to create a culture of cooperation and coordination between health, social care, public health, other local

⁵⁰⁷ Competition and Markets Authority, “Consumer Protection: Guidance to the CMA’s Approach to Use of Its Consumer Powers,” March 2014, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/288624/CMA7_Consumer_Protection_guidance.pdf; Department for Business, Innovation and Skills, “Empowering and Protecting Consumers: Government Response to the Consultation on Institutional Reform,” April 2012, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/253701/bis-12-510-empowering-protecting-consumers-government-response-1.pdf.

⁵⁰⁸ Competition and Markets Authority, “Consumer Protection: Guidance to the CMA’s Approach to Use of Its Consumer Powers,” 12.

services and the third sector. Working in silos is no longer acceptable.”⁵⁰⁹ The MHRA is currently neither a partner nor even mentioned in the context of these associations, despite being an important – but currently ‘siloed’ – partner in both of these goals. Collaborating with these associations will help ensure that citizens’ medical *and* non-medical interests form the basis of arm’s length licensing decisions.

⁵⁰⁹ National Collaboration for Integrated Care and Support, “Integrated Care and Support: Our Shared Commitment,” May 2013, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/287815/DEFINITIVE_FINAL_VERSION_Integrated_Care_and_Support_-_Our_Shared_Commitment_2013-05-13.pdf, 1.

Conclusions

In this concluding chapter, I highlight the original contributions that this dissertation makes to both the immediate practice of arm's length governance and the more theoretical debates in the literature on public reasoning and deliberation, ethics expertise, and patient and public involvement. First, I list substantive recommendations for both the Government in creating and managing arm's length bodies and the HFEA, NICE, and MHRA as individual arm's length bodies. I also suggest the broader applications of public accountability as reciprocity for policymaking on emerging biotechnologies, both within the UK and internationally. Second, I illustrate how the analytic and normative work undertaken in constructing this framework addresses a number of limitations, gaps, and confusions in the current bioethical and applied political philosophical literature.

1. Recommendations for Government: endorsing public accountability as reciprocity

My primary recommendation for policymakers in the Department of Health, the Cabinet Office, and Parliament is to endorse my framework for public accountability for those arm's length bodies that regulate emerging biotechnologies. While I have considered the HFEA, NICE, and MHRA here, similar issues arise in a number of other arm's length settings, such as NHS England, NHS Blood and Transplant, Public Health England, the Human Tissue Authority, the Care Quality Commission, and the Health Research Authority. While further inquiry is necessary to determine whether and how

public accountability as reciprocity would apply in these bodies, their shared politico-institutional form and decision-making about scientifically complex and morally controversial innovation in medicine suggests that it may be a promising approach.

The four conditions of my framework, its substantive commitment to reciprocal reasoning, and its procedural implications for decision-making could be endorsed as broad guidance in the form of a policy paper⁵¹⁰ from the Department of Health to its arm's length bodies, added into the accountability provisions of the yearly framework agreements between the Department and its non-departmental public bodies like the HFEA⁵¹¹ and NICE,⁵¹² or introduced as recommendations in the Department's triennial reviews of each arm's length body.

Recommendation 1. The Government should adopt the four conditions of public accountability as reciprocity, highlighting their procedural implications for deliberation and decision-making within arm's length bodies:

1. *Reasoning reciprocally:* Given the conflicting but reasonable sets of values and beliefs that the UK public holds with regards to novel biotechnologies, bodies should publicly adopt a reciprocal method of reasoning aimed at identifying mutually justifiable policy positions: ones based on reasons that all persons who recognize others as free and equal could accept. This means abandoning both attempts to locate the moral truth of the matter and attempts to 'democratically' represent the views of the public.
 - In the appointment of members: all members must be willing to have their minds changed by argument and evidence in line with the demands of reciprocity. Members should come from diverse backgrounds, including scientific and clinical experts and non-expert 'lay' members of the public. The inclusion of a philosopher, or other similarly skilled person, capable of facilitating the process of reciprocal reasoning is desirable, although she is not afforded any special authority in

⁵¹⁰ See, for an example of similar guidance, Department of Health, "Liberating the NHS: Report of the Arm's-Length Body Review."

⁵¹¹ "Framework Agreement between the Department of Health and the Human Fertilisation and Embryology Authority," 2014, [http://www.hfea.gov.uk/docs/HFEA_Framework_Agreement_v16_\(FINAL_signed_for_publication\).pdf](http://www.hfea.gov.uk/docs/HFEA_Framework_Agreement_v16_(FINAL_signed_for_publication).pdf).

⁵¹² "Framework Agreement between the Department of Health and the National Institute for Health and Care Excellence," 2014, <https://www.nice.org.uk/Media/Default/About/Who-we-are/NICE-and-Dept-of-Health-Framework-Agreement-March-2014.pdf>.

determining which substantive reasons or arguments are properly reciprocal.

- In the consultation of the public: public consultation for the purpose of identifying mutually acceptable reasons or values (distinct from secondary forms of consultation geared towards identifying effective ways of meeting predetermined goals) must be tailored to elicit the range of values that citizens can recognize as relevant to the public decisions facing the body. Deliberative focus groups or standing citizens' councils are preferable to opinion polls or questionnaires for this purpose. Results feed into the body's attempt to locate shared values, but members must scrutinize these results to screen out positions that do not properly recognise other citizens as free and equal.
2. *Accepting deliberative disagreement*: Questions of both the permissibility and public rationing of emerging biotechnologies, due to special characteristics of these techniques or products, may predictably be ones to which reciprocal reasoning is incapable of yielding a definitive policy solution: bodies must accept that there may be multiple positions supported by different weightings of reciprocal reasons or that there may be no positions that do not require invoking non-reciprocal reasons.
 - In decision-making procedures: bodies must publicly recognize the existence of deliberative disagreement where it persists after sincere efforts to reason reciprocally. This implicates a strong form of transparency, in which bodies must flesh out such residual disagreement in publicly available and accessible accounts of meetings and decisions, and diligently avoid the façade of false consensus.
 3. *Voting*: Bodies should attempt a moral compromise in the face of deliberative disagreement. However, if such a compromise is not reached, bodies may legitimately base decisions on a majority vote in which members cast their votes on the basis of what they find to be the most convincing arguments after the process of reciprocal reasoning.
 - In decision-making procedures: bodies must publish the results of this voting process, which may include majority and minority statements, in the public document explaining the final decision.
 4. *Fostering the means of resistance*: Because decisions made in the face of deliberative disagreement will be based upon reasons that some citizens may reasonably reject, bodies have a duty to foster the means by which citizens may deliberatively contest these outcomes.
 - In institutional avenues for challenge and revision: bodies must create an appeals process and/or publicly announce at the time a decision is released when it will be considered for revision. Reports of decisions and explanations of the decision-making process must provide an account sufficient for citizens to deliberatively engage within these avenues.

- In public engagement: bodies must maintain public engagement programmes that educate citizens about both the role of reciprocity in making arm's length public decisions and the facts and uses of the emerging biotechnologies that they regulate.

A. International applications

While I have focused solely on the UK context, a number of other countries maintain similar arm's length bodies that regulate public access to emerging biotechnologies. These bodies utilize many of the features of decision-making addressed here, including public consultation, patient involvement, lay membership, and philosophical expertise. For example, the Australian National Health and Medical Research Council (NHMRC) is an independent statutory agency responsible for funding research and developing guidelines for both clinical practice and health research.⁵¹³ The Australian Health Ethics Committee, one of its statutory committees, is responsible for developing guidelines for human research and advising the Council on ethical aspects of health, for which it “undertakes formal public consultation when developing guidelines” and seats members with “expertise in philosophy, the ethics of medical research, public health and social science research, clinical medical practice and nursing, disability, law, religion, and health consumer issues.”⁵¹⁴ It is currently carrying out a policy review and corresponding public consultation on the clinical use of reproductive technologies, including whether to relax current guidelines on non-medical sex selection.⁵¹⁵ The guidelines produced by the Committee are used in the accreditation process for fertility

⁵¹³ National Health and Medical Research Council, “About the NHMRC,” last updated May 21, 2015, <https://www.nhmrc.gov.au/about>.

⁵¹⁴ National Health and Medical Research Council, “Australian Health Ethics Committee (AHEC),” last updated July 1, 2015, <https://www.nhmrc.gov.au/about/nhmrc-committees/australian-health-ethics-committee-ahec>.

⁵¹⁵ National Health and Medical Research Council, “An Invitation to Make a Submission: Review of Part B of the Ethical Guidelines on the Use of Assisted Reproductive Technology in Clinical Practice and Research,” July 15, 2015, <http://consultations.nhmrc.gov.au/files/consultations/drafts/artdraftethicalguidelines150722.pdf>.

clinics, and as such greatly influence public access to new techniques. In its political form and remit, then, it is similar to UK arm's length bodies in general and the HFEA in particular, and may find these recommendations useful for reforming or clarifying the justifications for its decision-making procedures.

Other international arm's length bodies include the German Ethics Council⁵¹⁶ and the French Biomedicine Agency,⁵¹⁷ as well as resource allocation bodies common to countries with public healthcare and licensing agencies common to technologically advanced democracies more generally. While not permanent arm's length bodies, many countries also rely upon one-off independent commissions of inquiry, like that of the Warnock Committee, to provide legislative or regulatory recommendations on particularly transformative and controversial developments in biomedicine. Historically, these tend to respond to developments in the area of assisted reproduction and embryo research, and examples include Canada's Baird Commission,⁵¹⁸ Australia's Waller Committee⁵¹⁹ and Lockhart Committee,⁵²⁰ and Spain's Special Commission for the Study of Human In Vitro Fertilisation and Artificial Insemination.⁵²¹

To the extent that such short-term commissions face the same legitimacy problem of independently identifying policy recommendations that are 'in the public interest,' they may benefit from the reciprocal approach to reasoning and public consultation

⁵¹⁶ Deutscher Ethikrat, "About Us: Ethics Council Act," <http://www.ethikrat.org/about-us/ethics-council-act>.

⁵¹⁷ Agence de la biomedecine, "About Us," <http://www.agence-biomedecine.fr/About-us?lang=fr>.

⁵¹⁸ Sonya Norris and Marlisa Tiedemann, "Legal Status at the Federal Level of Assisted Human Reproduction in Canada," Library of Parliament Background Paper (Ottawa: Library of Parliament, April 9, 2015).

⁵¹⁹ Louis Waller, "Reproductive Technology: In Australia, the Debate Moves to Embryo Experimentation," *The Hastings Center Report* 17, no. 3 (1987).

⁵²⁰ National Health and Medical Research Council, "Review of Human Cloning and Human Embryo Research Legislation (2005)," last updated July 17, 2015, <https://www.nhmrc.gov.au/about/nhmrc-committees/embryo-research-licensing-committee/human-embryos-and-cloning/review-human-cl>.

⁵²¹ For this and a dated but extensive list of other such commissions, see LeRoy Walters, "Ethics and New Reproductive Technologies: An International Review of Committee Statements," *The Hastings Center Report* 17, no. 3 (1987).

provided by my framework of public accountability. Indeed, an account of the Lockhart Committee's approach to decision-making by former members notes that it purposefully did not endorse any substantive principles or a specific ethical approach, preferring to ground its work in "a commitment to shared moral values, to a sophisticated understanding of community, and to democratic moral discourse factors," as well as "broad community consultation."⁵²² However, they note that some may nevertheless criticise their approach for being too 'messy.' My framework, based upon similar commitments, may provide some 'order' to and promote the replicability of this approach.

The United States Food and Drug Administration, in its licensing of medicines and medical devices, is set up at arm's length and serves a role analogous to the MHRA. It faces many of the same borderline problems encountered by the MHRA in regulating new devices, and many of my claims in Chapter VI about reconceptualising the distinction between the medical and the non-medical in accounting for risk and benefits pertain to its work, as I have suggested elsewhere.⁵²³ But such arm's length regulation of emerging biotechnologies is otherwise relatively rare in the US. Apart from questions about the federal funding of stem cell research, which has remained largely tied up with abortion politics in America, the US has historically avoided national regulation of access to assisted reproductive technologies, preferring to leave such questions to self-regulation by the medical community and the states. Furthermore, as it does not maintain a large scale, publicly-funded and -operated healthcare system like the UK, it has no technology appraisal programme analogous to NICE; indeed, the representation of such rationing schemes as tantamount to 'death panels' in the (factually misplaced) political

⁵²² Loane Skene et al., "The Lockhart Committee: Developing Policy through Commitment to Moral Values, Community and Democratic Processes," *Journal of Law and Medicine* 16, no. 1 (2008): 138.

⁵²³ Kyle T. Edwards and Caroline J. Huang, "Bridging the Consumer-Medical Divide: How to Regulate Direct-to-Consumer Genetic Testing," *Hastings Center Report* 44, no. 3 (2014).

opposition to the *Patient Protection and Affordable Care Act of 2010* suggests one type of obstacle that public rationing would face in the US.

Nevertheless, I suspect that 1) emerging developments in human genetic engineering, like CRISPR,⁵²⁴ that reinvigorate political and public debate about ‘designer babies,’ and 2) the rising cost of healthcare coupled with the aging population may transform American’s political approach to emerging biotechnologies within the next decade (perhaps two). In calling for a moratorium on the use of CRISPR to edit the human genome, US scientists and bioethicists have highlighted the need for “a discussion that bridges the research community, relevant industries, medical centers, regulatory bodies, and the public to explore responsible uses of this technology,”⁵²⁵ and have advocated future consideration of its applications using “some combination of expert and public involvement and debate, politically open and legitimate decision-making, and controlled implementation.”⁵²⁶ Similarly, the Affordable Care Act created the controversial Independent Payment Advisory Board (subject of the ‘death panel’ accusations) charged with “reducing the per capita rate of growth in Medicare spending.”⁵²⁷ While explicitly prohibited by the Act from recommendations to ‘ration care,’ it will be given in 2020 “additional tools like the ability to consider value-based benefit design and policies.”⁵²⁸ Despite America’s anti-rationing political rhetoric, the country already implicitly rations care⁵²⁹ and it seems likely that this Board will need to

⁵²⁴ Jeffrey D. Sander and J. Keith Joung, “CRISPR-Cas Systems for Editing, Regulating and Targeting Genomes,” *Nature Biotechnology* 32, no. 4 (2014).

⁵²⁵ David Baltimore et al., “A Prudent Path Forward for Genomic Engineering and Germline Gene Modification,” *Science* 348, no. 6230 (2015).

⁵²⁶ Hank Greely, “Of Science, CRISPR-Cas9, and Asilomar,” *Law and Biosciences Blog*, April 4, 2015, <https://blogs.law.stanford.edu/lawandbiosciences/2015/04/04/of-science-crispr-cas9-and-asilomar/>.

⁵²⁷ Jim Hahn and Christopher M. Davis, “The Independent Payment Advisory Board,” CRS Report for Congress (Washington, DC: Congressional Research Service, April 17, 2013): ii.

⁵²⁸ *Ibid.*, 29.

⁵²⁹ U. E. Reinhardt, “Rationing Health Care: What It Is, What It Is Not, and Why We Cannot Avoid It,” *The Baxter Health Policy Review* 2 (1996).

explicitly confront questions of public resource allocation if attempts to cut costs by reducing inefficiencies in the provision of care are insufficient.⁵³⁰

If such predictions come to pass, my framework offers a path forward for a society relatively new to the arm's length regulation of public access to emerging biotechnologies, one that draws on the experiences of the country arguably most familiar with it. Particularly given the extent of fundamental religious and moral disagreement in the US, a reciprocal approach coupled with process-oriented conditions to manage the persistence of deliberative disagreement is crucial to the legitimacy of decisions on assisted reproduction and resource allocation.

All the examples considered in this section come from governments that may be broadly characterized as 'liberal democratic,' in which the power of popularly-elected representatives is checked by a set of either written or unwritten individual rights and liberties, and thus are politically similar to the UK. Nevertheless, any attempt to export this set of claims about public accountability as reciprocity from the UK to other countries must carefully consider the socio-cultural variation within this group of liberal democratic states that may affect the application and viability of the approach. For instance, it would likely face significant political opposition in the US, and so extensive and popularly visible public engagement would be needed prior to any attempts to formally institute it.

⁵³⁰ As early as 1990, US healthcare scholars argued that "eliminating inefficiencies in the system can provide brief fiscal relief, but rationing of beneficial services, even to the well-insured, offers the only prospect for sustained reduction in the growth of health care spending." See H. Aaron and W. B. Schwartz, "Rationing Health Care: The Choice Before Us," *Science* 247, no. 4941 (1990).

2. Recommendations for arm's length bodies: applying public accountability as reciprocity

An important observation from the three 'case study' chapters, which examine how the theoretical framework may be applied towards the practical resolution of decision-making challenges, is that the endorsement of the four conditions is merely a first step in institutional attempts to improve the legitimacy of public decisions. This is particularly clear in the case of NICE, which – while already endorsing accountability for reasonableness, an approach similar in many ways to my own – nevertheless has struggled to locate an appropriate role for the patient voice in technology appraisals. Applying public accountability as reciprocity and understanding what follows for the structure of processes and stakeholder roles requires assiduous attention to institutional context and the nature of the decisions for which the body is responsible.

A. Recommendations for the HFEA

In the chapter on the HFEA, I argued that the revealed method of 'balancing' considerations from the general public and various stakeholders fails to make transparent and publicly account for the way in which this method of reasoning gives weight to certain considerations over others. Importantly, it is possible to recognize the shortcomings of this method as a process of public justification however much one may agree that reasons having to do with health and welfare – those apparently embedded within the HFEA's method of balancing – are the right substantive ones to prioritize. In its place, I proposed a reciprocal method of reasoning, in which the 'weight' afforded to various considerations or arguments is based on whether the values underpinning them are mutually justifiable to citizens who recognize each other as free and equal. This method of reasoning is candid about the substantive motivation or goal – reciprocity –

underlying its approach to public decision-making. It can also adequately respond to challenges that the role of public bodies making value judgments on behalf of society should be either 1) to reflect public views and values, or 2) to identify the policy position most capable of withstanding philosophical debate, all things considered.

Recommendation 2. The HFEA should adopt a reciprocal method of reasoning, in which the substantive commitment to mutual acceptability drives the evaluation of reasons and arguments in deliberation. This commitment must be made publicly rather than represented as a purely procedural (and falsely neutral) approach to decision-making.

Committing to this reciprocal method of reasoning has key implications for the HFEA's decision-making procedures, in particular its method for consulting the public and the justifiable scope of the philosopher's role as a member.

Recommendation 3. Because reciprocal reasoning does not seek to uncritically reflect public views or values, the HFEA should discontinue representative polling of the public in its policy reviews, relying instead upon focus groups or a citizens' council.

Recommendation 4. The HFEA should maintain a philosopher-member on its committee to help guide the process of reciprocal reasoning, but she shall have no special authority over lay members and scientific members in determining which public values or reasons are properly reciprocal.

Like other government bodies, the HFEA has a legal duty to publicly consult when considering significant changes to its policies or services. The viability of transitioning to a more deliberative approach to public consultation, using focus groups or a citizens' council like that of NICE, is thus partially dependent upon superior bodies, including the Department of Health and the Administrative Court, recognizing the necessity and sufficiency of this kind of public consultation for the decisions that the HFEA makes. Of course, wider public engagement around these smaller scale methods of consultation is vital to the HFEA's transparency about the nature of its decisions.

B. Recommendations for NICE

NICE has already adopted a method of reasoning that is largely in line with my account of reciprocity: accountability for reasonableness. It properly recognizes that ‘relevant’ reasons or values for resource allocation – those set out in *Social Value Judgements* – should be determined by reference to the values of the general public, rather than solely to those of patient interest groups or pharmaceutical companies. While further scrutiny of this stage of decision-making is warranted, public consultation through NICE’s Citizens Council followed by reciprocal reasoning by the NICE Board are *prima facie* suitable procedural requirements to meet this goal. Instead of addressing this stage of institutional decision-making (one relatively analogous to that examined in the HFEA context), I examined the stage of technology appraisals. Here, Technology Appraisal Committees must determine what the substantiation of these pre-determined social value judgements requires in the context of a specific decision on whether to fund a novel and expensive treatment on the NHS. Within this context, NICE has struggled to clarify a role for the patient voice in bearing out these values.

Recommendation 5. NICE should recognize and publicly affirm that the role of patients as consultees in technology appraisals is limited to evidence provision, and should adjust its patient involvement processes accordingly. In particular, the ‘patient expert’ is a poor method for sourcing reliable and representative information capable of addressing the limitations of RCTs and cost effectiveness modelling. Emphasis should instead be placed upon facilitating evidence submissions from patient groups, which can then leverage organizational resources towards the systematic collection of information from patient-members.

Patients with conditions that are rare or who are poorly organised due to cultural or socio-economic factors may deserve additional resources and guidance in gathering and presenting this kind of evidence. Additionally, patients may still valuably serve as lay members on Appraisal Committees or Appeal Panels, on justifications set out in the HFEA context.

Once NICE recognizes the need to treat patients as evidence providers in technology appraisals, and reforms its involvement mechanisms to create viable avenues for this input, Appraisal Committees are faced with a second practical challenge: how to interpret and weigh this kind of evidence against clinical and cost effectiveness data. Specifically, NICE must be able to manage disagreement in a publicly accountable manner about what this patient evidence (as part of the broader evidence base) means for the substantiation of its social value judgements. Although NICE has an appeals process in which groups may challenge a decision on the ground that it is ‘unreasonable’ in light of the evidence, its accountability for reasonableness approach only requires the Appeal Panel to confirm that the decision was based on a reasonable interpretation of the evidence – even if other reasonable interpretations supporting conflicting recommendations existed. In such cases of reasonable disagreement, my public accountability framework requires NICE to actively support the means by which those who ‘lose out’ in such resolutions may deliberatively engage with and contest the outcome in the future.

Recommendation 6. NICE should create a formal, institutional feedback loop from the Appeal Panel to the Evidence Review Group and finally back to the Public Involvement Programme. The Panel would be responsible for identifying during the appeals process instances in which there was *reasonable* disagreement about the interpretation of evidence. The Evidence Review Group would consider these cases and clarify evidential expectations based upon current, methodological best practice: how patient groups would need to collect and present information for it to be considered more reliable within the context of cost effectiveness analyses. It would also pass these explanations on to the Public Involvement Programme, which would in turn use them to adjust evidence submission templates and guidance.

This feedback loop appropriately fosters the means of deliberative resistance, the most unique condition of my public accountability framework. It provides specific patient groups with the kind of explanation needed to engage with NICE’s process when the treatment in question is reconsidered or new treatments for their condition are

scheduled for appraisal. Furthermore, it commits NICE to iteratively improving its method of patient involvement to the benefit of all patients and patient groups that engage in its work; this may be expected to enhance the quality of patient evidence NICE receives and thereby contribute to the broader NHS goal of patient-centred medicine.

C. Recommendations for the MHRA

The MHRA initially presents as an arm's length body responsible solely for fact-based, technical judgments that support the protection and promotion of public health; unlike NICE and the HFEA, it neither presents itself nor is commonly recognized as a body making *value* judgements to which patients and the public will be held in accessing medicines and medical devices. In diagnosing the borderline problems posed by a growing number of novel devices, the MHRA has thus naturally fallen back upon the traditional solution of better and more appropriate technical expertise, aiming to strengthen its access to clinical expertise in the area of medical devices. I illustrated that the root of these borderline problems can be traced to uncertainty about whether and how to subject these new devices to the MHRA's traditional assessments of risk-benefit and effectiveness. By considering tDCS devices and telehealth devices like fall detectors and GPS trackers, I argued 1) that these assessments entail significant judgements about values, and 2) that the values 'relevant' to patients and the public in making these assessments – the parties bound by these licensing decisions – are far broader than the 'medical' or 'health-related' considerations to which the MHRA traditionally limits itself.

So, while the protection of public health necessitates MHRA assessment of the health risks of non-therapeutic cognitive enhancement devices, the acceptability of these

risks can only be *reasonably* considered against the devices' non-medical benefits, which are likely to be valued differently by various sectors of the public. Similarly, certain telehealth devices pose risks that do not strictly affect patient health but may fundamentally affect other components of patients' well-being, such as autonomy and privacy, and thus are relevant to risk-benefit assessment and risk management. Perceiving the true nature of the MHRA's risk-benefit and effectiveness analyses reveals that the borderline problems caused by emerging biotechnologies cannot be solved simply by bolstering technical expertise. Rather, the MHRA must engage in public deliberation with potential users about what constitutes a reasonable trade-off between risks and benefits and what values or factors are relevant to these concepts.

Recommendation 7. The MHRA should open up its value judgments about what constitutes a reasonable trade off of risks and benefits – as well as what factors count in measuring risks and benefits – to a process of public deliberation. In particular, the MHRA should consult with potential users at two key stages: the design of clinical trials and the post-trial assessment of risks and benefits. This will require relaxing its strict distinction between the 'medical' and the 'non-medical'/'consumer' in order to account for the range of considerations relevant to citizens in these public decisions to grant or limit access.

In recognizing this wider range of considerations relevant to its decision-making, the MHRA will also need to construct functional bridges with those bodies with whom it shares regulatory boundaries, in order to ensure that neither important risks nor benefits of emerging biotechnologies fall through the gaps between institutions. The MHRA has remained uniquely siloed from both health and social care and consumer collaborations alike, a state of affairs that must be rectified.

Recommendation 8. The MHRA should create stronger, formal connections to other health care, social care, and consumer protection bodies, in particular the Care Quality Commission and Local Trading Standards Authorities. This may be partially accomplished by joining existing collaborations, like the Consumer Protection Partnership and the National Collaboration for Integrated Care and Support.

Both of these recommendations will require extensive coordination with the European Commission, given the legal precedence of its medical directives and regulatory instruments. Although it may take time, as a major leader in medical regulation at the European level, the MHRA should be capable of advocating on behalf of this more holistic and publicly accountable approach to considering the way novel medicines and devices impact citizens' well-being.

3. Contribution to the practical application of public reason and deliberative democracy

Accounts of public reason and deliberative democracy have proved extremely influential in political philosophy over the past three decades, and are two of the most important recent contributions to the theory of democratic legitimacy. Nevertheless, the translation of their theoretical claims into practice is often hindered by the abstract or vague nature of driving concepts – like ‘reasonableness,’ ‘mutual acceptability,’ and the ‘ideal speech situation.’ My framework for public accountability in the arm’s length context contributes to a range of recent efforts to apply these theories to practical, political decision-making in order to realize the benefits that they promise to the quality and legitimacy of public decisions.

The work of Norman Daniels and James Sabin is perhaps the most well known attempt in this endeavour, as well as one frequently taken up by actual policymakers in structuring decision-making. This latter point suggests that in addition to any normative legitimacy secured by accountability for reasonableness, it may also claim a substantial degree of empirical legitimacy – that is, people *believe* that it is a normatively legitimate approach to policymaking, as well as one that can be practically implemented. Because

of these three strengths of their work – its claim to normative legitimacy, empirical legitimacy, and practicability – it is important to understand how my framework of public accountability as reciprocity is both distinct from and an improvement upon accountability for reasonableness.

First, and most significantly, my framework is a more theoretically rigorous interpretation of the demands of normative legitimacy for public decisions made in the face of reasonable disagreement. Accountability for reasonableness is offered primarily as a fair, publicly acceptable *process* for healthcare rationing, given widespread disagreement – both philosophical and empirical – about the substantive principles that should govern priority setting. And indeed three of its four conditions are truly procedural in nature and widely accepted: publicity, challenge and revision, and enforcement. However, its condition of relevance – that decisions must be based upon reasons and evidence that ‘fair-minded’ people recognize as relevant - imports constraints on the *substance* of reasons and evidence that carry justificatory force, constraints that may be more or less restrictive depending upon how narrowly we define the group of fair-minded people.

Reciprocity, my framework’s analogue to Daniels and Sabin’s relevancy, does not escape this substantive constraint but it candidly recognises its centrality in driving other procedural requirements, which I take to be an important component of public transparency. Furthermore, it takes seriously the possibility of persistent deliberative disagreement, and the significance of having to make politically coercive decisions in the face of such disagreement. Whereas accountability for reasonableness depends upon there being broad agreement amongst fair-minded people about the reasons and evidence ‘relevant’ to a decision, my framework recognizes that, in practice, a group of people all of whom appear to recognize others as free and equal and seek fair terms of cooperation

may still disagree about what reasons are relevant to the decision at hand. It recognizes the potential *indeterminacy* of such attempts at public reasoning, understood as cases in which none of the policy options can avoid invoking reasons or evidence the relevance of which is disputed, which I argued may often be the case in decisions about reproductive technologies that implicate a position on the moral or metaphysical status of the embryo or foetus.

I also suggested that we may be more likely to agree on the relevancy of reasons in the allocation of scarce healthcare resources, suggesting that this criticism of accountability for reasonableness is not as devastating when it is applied in this context. However, Daniels and Sabin do recognize what I have explained as the potential *inconclusiveness* of public reasoning even in cases where we agree on the set of relevant reasons and evidence. This relates to deliberative disagreement, caused by characteristics similar to Rawls' burdens of judgement, about how to order or weight a set of relevant values or how to interpret recognisably relevant evidence in basing decisions upon these values. In such cases, accountability for reasonableness recognises the outcome of a final vote as legitimate because all agreed on the relevancy of the reasons underlying it, even if deliberators in the minority supported a conflicting position based on a reasonable weighing of the same values or evidence.

My framework properly takes both inconclusiveness and indeterminacy more seriously than accountability for reasonableness, by recognising and accounting for the significance of coercing citizens on the basis of reasons that fair-minded people may either reject or may interpret as supporting a conflicting decision. While recognizing the necessity of making decisions within a reasonable timeframe even under these conditions, my framework requires a step 'beyond' accountability for reasonableness, that of actively fostering the means by which citizens can engage with the grounds for

decisions and contest them deliberatively (through reciprocal reasoning) in the future. This underpins a requirement for candour about the existence of alternative reasonable options, or the moral and metaphysical disagreements that prevent decisions from being made with reference to exclusively reciprocal reasons. It also requires public education about reciprocity in addition to the developments in science and technology that a body regulates. As my case study of NICE illustrated, while avenues for appeals or revision are necessary to meet this condition, they are not sufficient.

In addition to better capturing the demands of normative legitimacy, my framework is also more practically useful. Accountability for reasonableness is tailored specifically to healthcare resource allocation decisions, whereas public accountability as reciprocity can be consistently endorsed as a method for making independent but accountable decisions across the wide range of bodies responsible for regulating emerging biotechnologies. Because it can properly address situations of indeterminacy that are likely to arise in the area of assisted reproduction and perhaps in future developments in the area of human enhancement, it can be applied in contexts beyond resource allocation in a way that accountability for reasonableness cannot. This is a key strength given the way in which funding decisions are fundamentally linked to and constrained by licensing decisions, and prescription and purchasing decisions are in turn dependent upon these two earlier stages. Having a common conception of accountability across these settings will promote coordinated and consistent approaches to novel treatments and products. Indeed, in addition to working through specific problems currently facing the HFEA, NICE and the MHRA, the three case study chapters serve more broadly to illustrate that my framework can be practically and fruitfully applied in such diverse policy settings.

Finally, I have attempted to illustrate what reciprocity demands for the reform of decision-making processes, laying out what I expect to be widely acceptable procedural requirements – like publicity, transparency, candour, and public engagement – that follow from a commitment to reciprocity, an acceptance of deliberative disagreement, and the fostering of the means of resistance. This is in contrast to an approach that seeks to identify the substantive principles on which decisions should be based. I take it that accountability for reasonableness’s practical appeal and claim to empirical legitimacy stems largely from its similar proposal of a process by which people can agree to abide. Although my account is more transparent about the substantive commitment motivating this process, I expect that this procedural orientation will similarly augment its empirical legitimacy. In connecting the identification of reciprocal reasons to a properly tailored form of public consultation, it also secures and justifies a process of public involvement, whereas, as Peter Littlejohns and Katharina Kieslich have recently argued, accountability for reasonableness’s “relation to efforts of public participation remains opaque.”⁵³¹

4. Implications for the concept of ‘ethics expertise’

The question of whether and what kind of expertise the practical ethicist or bioethicist possesses is the subject of a fairly entrenched debate, invoking a variety of metaphysical and epistemological claims about the existence of moral truth and the ability of the ethicist to better discern these truths than the average person. Peter Singer is perhaps the most well known supporter of the moral philosopher as moral expert, one who is, “in general, better suited to arrive at the right, or soundly based, moral

⁵³¹ Katharina Kieslich et al., “Does Accountability for Reasonableness Work? A Protocol for a Mixed Methods Study Using an Audit Tool to Evaluate the Decision-Making of Clinical Commissioning Groups in England,” *BMJ* 5, no. 7 (2015).

conclusions than non-philosophers.”⁵³² Representing a kind of middle ground, Lisa Rasmussen distinguishes between this ‘strong’ account of the ethicist’s expertise, which she disavows, and a ‘soft’ account that she endorses as “the ability to give non-normatively binding recommendations grounded in a pervasive ethos or practice within a particular context.”⁵³³ Finally, the work of Christopher Cowley falls at the far end of the spectrum, arguing not only that ethicists are not moral experts but also that there is no such thing as moral expertise because “[m]oral reasons and judgements are always declared within an embodied moral perspective on the situation...not there in the situation, to be discovered and applied by whomever stumbles across them.”⁵³⁴

My work does not set out a new account of ethics expertise to contend with the range of accounts from which these three are drawn. Rather, I argue that *whatever* knowledge or skills we think the ethicist possesses, her role and authority in public policymaking is contingent upon the demands of reciprocity. As such, for our practical purposes, we can largely sidestep this ongoing debate about the nature of ethics expertise in determining the role for the ethicist in making collectively binding decisions.⁵³⁵ As discussed in Chapter IV in relation to the HFEA, even if we think that ethicists possess the strong version of moral expertise, what follows from this recognition is quite limited: the inference from this expertise to any kind of political authority, what David Estlund calls the expert/boss fallacy, is blocked by the theoretically prior claims of reciprocity about what we owe other citizens in collective decision-making once we recognize them as free and equal. In exercising the power of the state, and granting the range of

⁵³² Peter Singer, “Moral Experts,” 17.

⁵³³ Lisa M. Rasmussen, “An Ethics Expertise for Clinical Ethics Consultation,” *The Journal of Law, Medicine & Ethics* 39, no. 4 (2011): 650.

⁵³⁴ Christopher Cowley, “A New Rejection of Moral Expertise,” 276.

⁵³⁵ Kyle T. Edwards and Michael Dunn, “Ethics Expertise in Public Policymaking,” *Journal of Medicine and Philosophy*, in forthcoming Special Issue on Ethics Expertise eds. Lisa Rasmussen and Ana Iltis, (expected 2016); Kyle T. Edwards and Zuzanna Deans, “Empirical Ethics and the Role of the Professional Ethicist in Public Policymaking,” in *Empirical Bioethics: Practical and Theoretical Perspectives*, ed. Jon Ives, Michael Dunn, and Alan Cribb, Expected (Cambridge: Cambridge University Press, expected 2015).

reasonable conceptions of the good life held within it, we must give our fellow citizens reasons for coercive decisions that they can, on some level, accept – even if we are convinced that we know the ‘right’ answer to such decisions.

As such, the ability (or lack thereof) of the ethicist to identify the right answer to moral questions in public policymaking is fairly immaterial. She shares the arm’s length bodies’ delegated authority equally with all other members and is thus an equal partner in seeking out mutually justifiable policy options. Because of her training and experience in constructing and critiquing arguments, she may be an ideal member to guide the process of reciprocal reasoning and give a public account of this process in the final policy document. However, she has no special authority to identify the substantive values that are properly reciprocal and how these values should be weighed against each other.

5. Implications for the Patient and Public Involvement movement

Finally, the claims that I have made about public accountability clarify the justification for patient and public involvement in the deliberately independent decision-making of arm’s length bodies, and suggest how the current practice of involvement must be reformed in light of this justification. The PPI movement in healthcare has been going strong in the UK for at least two decades,⁵³⁶ but its rationale is rarely made explicit.⁵³⁷ While many have previously suggested the importance of distinguishing between public involvement, on the one hand, and patient involvement, on the other, my work contributes an account of how to make this distinction within the policymaking context, one relatively overlooked compared to the medical research context. In the case

⁵³⁶ Baggott, “A Funny Thing Happened on the Way to the Forum?”

⁵³⁷ Florin and Dixon, “Public Involvement in Health Care.”

of NICE, for instance, public involvement is justified as a component in the process of identifying reciprocal (or, in its terms of accountability for reasonableness, relevant) values upon which to base resource allocation decisions. Specific patient groups interested in a decision about a particular treatment for their condition have no special role in setting these public value judgements, but a separate form of patient involvement is justified by their ability to provide unique evidence about the condition and technology in question.

Furthermore, the method of involvement must be tailored to its justification. In the HFEA, for example, eliciting the range of reasons that individuals recognize as relevant to a decision demands a deliberative space in which citizens can learn about, discuss, and debate the issues, such as a focus group or citizens' council; it will be difficult if not impossible with public consultation through polls or questionnaires that provide a large sample size but a shallow and simplistic account of public views and values. Including members of the general public as 'lay' members on decision-making committees must be recognized as distinct from the practice of public consultation; lay members enhance reciprocal reasoning, due to their lack of 'disciplinary blinkers' with which expert-members may unwittingly approach questions, but they cannot represent the public or identify public values in the way that consultation does. This moves us beyond justifications for lay membership based on either their lack of knowledge about an issue or their disinterestedness in it, usefully quelling worries about whether lay members become too 'professionalized' by virtue of serving on an arm's length body or whether they are truly disinterested if they sign on as members to a time-consuming and low-compensation role.

In terms of patient involvement, making judgements about risks and benefits (as the MHRA does) or costs and benefits (as NICE does) at the level of a large patient

population or even a patient sub-group requires representative and robust evidence from patients about the factors that affect quality of life, or those that count as risks and benefits. As such, non-representative methods of collecting evidence, such as patient experts giving oral evidence, will not suffice; patient involvement must be structured to elicit evidence submissions that will be recognized as robust when considered alongside the rest of the evidence base. This will require guiding patient groups to leverage their resources towards the systematic collection and analysis of patient evidence.

Proper patient involvement is not just about making a legitimate licensing or funding decision. These steps of regulating public access to emerging biotechnology are ‘upstream’ of the doctor-patient discussions that ultimately impact individuals’ health and broader well-being. If we want to realise patient-centred care in the clinic, we need to utilise patients as providers of robust evidence at both the licensing stage of the MHRA and the funding stage of NICE.

6. Future directions

Scientific and technological innovation in the health and life sciences offers considerable individual and population benefits, but continues to put pressure on our ability to agree on the structures, rules, and limits of collective life. As such, we must match biomedical innovation with innovation in the political theory and practice of legitimate public decision-making on access to emerging biotechnologies. Rather than seeking to avoid or deny the extensive value pluralism within UK society, we must actively engage with it through public deliberation, and embed reciprocity within our approach to collective decisions. My framework of public accountability provides some original and progressive developments as part of this larger endeavour. However, as the application of this framework to the diverse settings of the HFEA, NICE, and the MHRA

has illustrated, sustained and careful attention to the demands of the four conditions on public reasoning within various practical contexts will be necessary to chart a reciprocal approach towards the innovations on the horizon.

Bibliography

- Aaron, H., and W. B. Schwartz. "Rationing Health Care: The Choice before Us." *Science* 247, no. 4941 (January 26, 1990): 418–22. doi:10.1126/science.2300804.
- Abelson, Julia, Pierre-Gerlier Forest, John Eyles, Patricia Smith, Elisabeth Martin, and Francois-Pierre Gauvin. "Deliberations about Deliberative Methods: Issues in the Design and Evaluation of Public Participation Processes." *Social Science & Medicine* 57, no. 2 (July 2003): 239–51.
- Abelson, Julia, Mita Giacomini, Pascale Lehoux, and Francois-Pierre Gauvin. "Bringing 'the Public' into Health Technology Assessment and Coverage Policy Decisions: From Principles to Practice." *Health Policy* 82, no. 1 (June 2007): 37–50. doi:10.1016/j.healthpol.2006.07.009.
- Ackerman, Bruce A. *Social Justice in the Liberal State*. New Haven: Yale University Press, 1981.
- Agbonlahor, Winnie. "Interview: Dr Ian Hudson, MHRA." *Civil Service World*, May 20, 2014. <http://www.civilserviceworld.com/articles/interview/interview-dr-ian-hudson-mhra-0>.
- Agence de la biomedecine. "About Us," n.d. <http://www.agence-biomedecine.fr/About-us?lang=fr>.
- Alan Pifer. "Letter: On Quasi-Public Organizations; Whence Came the Quango, and Why." *The New York Times*, September 5, 1987, sec. Opinion. <http://www.nytimes.com/1987/09/05/opinion/l-letter-on-quasi-public-organizations-whence-came-the-quango-and-why-969587.html>.

- . “The Quasi Nongovernmental Organization.” Annual Report for 1967. Carnegie Corporation of New York, 1967. <http://files.eric.ed.gov/fulltext/ED018846.pdf>.
- Alfred Moore. “Between Competence and Consent: Democratic Theory and Expertise.” In *Expertise and Democracy*, edited by Cathrine Holst, Vol. 1. ARENA Report. Oslo: ARENA, 2014.
- . “Questioning Deference: Expert Authority in a Deliberative System.” Presented at the Annual Meeting of the Midwest Political Science Association, Chicago, 2011.
- Alzheimer’s Society. “Alzheimer’s Society Response to NICE Appraisal Consultation Document 2,” February 2006. <https://www.nice.org.uk/guidance/ta111/documents/alzheimers-consultee-and-commentator-comments-alzheimers-society2>.
- Amis, Lizzie. “Patient Involvement in NICE Technology Appraisals.” In *Patients, the Public, and Priorities in Healthcare*, edited by Michael Rawlins and Peter Littlejohns. Oxford: Radcliffe Publishing, 2009.
- Amis, Lizzie, and Heidi Livingstone. “Technology Appraisal Patient Expert Survey 2012 Report.” NICE Public Involvement Programme, April 2014. <http://www.nice.org.uk/media/default/About/NICE-Communities/Public-involvement/Public-involvement-programme/Patient-expert-TA-report-final-1.pdf>.
- Archard, David. “Why Moral Philosophers Are Not and Should Not Be Moral Experts.” *Bioethics* 25, no. 3 (2011): 119–27. doi:10.1111/j.1467-8519.2009.01748.x.
- Arnstein, Sherry R. “A Ladder Of Citizen Participation.” *Journal of the American Institute of Planners* 35, no. 4 (1969): 216–24. doi:10.1080/01944366908977225.
- Audi, Robert. *Religious Commitment and Secular Reason*. Cambridge: Cambridge University Press, 2000.

- Baggott, Rob. "A Funny Thing Happened on the Way to the Forum? Reforming Patient and Public Involvement in the NHS in England." *Public Administration* 83, no. 3 (2005): 533–51. doi:10.1111/j.0033-3298.2005.00461.x.
- Baltimore, David, Paul Berg, Michael Botchan, Dana Carroll, R. Alta Charo, George Church, Jacob E. Corn, et al. "A Prudent Path Forward for Genomic Engineering and Germline Gene Modification." *Science* 348, no. 6230 (April 3, 2015): 36–38. doi:10.1126/science.aab1028.
- Barro, Robert J. "The Control of Politicians: An Economic Model." *Public Choice* 14, no. 1 (March 1973): 19–42. doi:10.1007/BF01718440.
- Benhabib, Seyla. "Toward a Deliberative Model of Democratic Legitimacy." In *Democracy and Difference: Contesting the Boundaries of the Political*. Cambridge: Cambridge University Press, 1996.
- Berry, Katy. "Authority Policy on the Use of PGD for Lower Penetrance Inherited Cancer Susceptibility Conditions," May 10, 2006. http://www.hfea.gov.uk/docs/PGD_-_choices_and_boundaries_May06.pdf.
- Black, Nick. "Evidence Based Policy: Proceed with Care." *BMJ: British Medical Journal* 323, no. 7307 (August 4, 2001): 275–79.
- . "Why We Need Observational Studies to Evaluate the Effectiveness of Health Care." *BMJ* 312, no. 7040 (May 11, 1996): 1215–18. doi:10.1136/bmj.312.7040.1215.
- Bohman, James. "Survey Article: The Coming of Age of Deliberative Democracy." *Journal of Political Philosophy* 6, no. 4 (1998): 400–425.
- Bohman, James, and William Rehg. "Introduction." In *Deliberative Democracy: Essays on Reason and Politics*, edited by James Bohman and William Rehg. London: MIT Press, 1997.

- Braun, Kathrin, and Susanne Schultz. “‘... a Certain Amount of Engineering Involved’: Constructing the Public in Participatory Governance Arrangements.” *Public Understanding of Science (Bristol, England)* 19, no. 4 (July 2010): 403–19.
- Bridges, John F. P., and Christopher Jones. “Patient-Based Health Technology Assessment: A Vision of the Future.” *International Journal of Technology Assessment in Health Care* 23, no. 1 (2007): 30–35. doi:10.1017/S0266462307051549.
- Brown, Mark B. “Survey Article: Citizen Panels and the Concept of Representation*.” *Journal of Political Philosophy* 14, no. 2 (June 1, 2006): 203–25. doi:10.1111/j.1467-9760.2006.00245.x.
- Business Companion. “Electrical Equipment: Essential Information for Suppliers of Electrical Equipment, Including CE Marking.” In-Depth Guide, n.d. <http://www.businesscompanion.info/en/quick-guides/product-safety/electrical-equipment#Satisfyingthegeneralsafetyrequirement>.
- Buxton, Martin, Louise Longworth, James Raftery, Mark Sculpher, and Adrian Towse. “Reforming the Cancer Drug Fund.” *The BMJ* 349 (November 28, 2014): g7276. doi:10.1136/bmj.g7276.
- Cabinet Office. “Categories of Public Bodies: A Guide for Departments,” December 2012. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/80075/Categories_of_public_bodies_Dec12.pdf.
- . *Executive Non-Departmental Public Bodies 1999* (Cm 46, 1999). Cited in Katharine Dommett, “Triennial Reviews and Public Body Reform in the UK.,” ESRC Shrinking the State Project, January 2014.
- . *Government Response to the Public Administration Select Committee Report. “Smaller Government: Shrinking the Quango State.”* (Cm 8044, March 2011).

- . “Guidance on Reviews of Non Departmental Public Bodies,” June 2011. http://www.civilservice.gov.uk/wp-content/uploads/2011/09/triennial-reviews-guidance-2011_tcm6-38900.pdf.
- . “Public Bodies 2014: Summary Data,” June 3, 2015. <https://www.gov.uk/government/publications/public-bodies-2014/public-bodies-2014-summary-data>.
- . “Public Bodies Reform,” November 10, 2014. <https://www.gov.uk/public-bodies-reform>.
- . “Using the Public Bodies Act 2011: A Guidance and Best Practice Note for Officials,” September 2013. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/243888/Using_the_Public_Bodies_Act_2011.pdf.
- Cameron, David. “People Power: Reforming Quangos.” Speech to Reform, July 6, 2009. <http://conservative-speeches.sayit.mysociety.org/speech/601336>.
- Carpini, Michael X. Delli, Fay Lomax Cook, and Lawrence R. Jacobs. “Public Deliberation, Discursive Participation, and Citizen Engagement: A Review of the Empirical Literature.” *Annual Review of Political Science* 7, no. 1 (2004): 315–44. doi:10.1146/annurev.polisci.7.121003.091630.
- Cartwright, Nancy. “A Philosopher’s View of the Long Road from RCTs to Effectiveness.” *The Lancet* 377, no. 9775 (April 2011): 1400–1401. doi:10.1016/S0140-6736(11)60563-1.
- Cartwright, Nancy, and Eileen Munro. “The Limitations of Randomized Controlled Trials in Predicting Effectiveness.” *Journal of Evaluation in Clinical Practice* 16, no. 2 (April 1, 2010): 260–66. doi:10.1111/j.1365-2753.2010.01382.x.
- Cerri, Karin H., Martin Knapp, José-Luis Fernández, and Jose-Luis Fernandez. “Decision Making by NICE: Examining the Influences of Evidence, Process and

- Context.” *Health Economics, Policy, and Law* 9, no. 2 (April 2014): 119–41. doi:10.1017/S1744133113000030.
- Chairman Lord Nolan. “Standards in Public Life: First Report of the Committee on Standards in Public Life.” Cm 2850-I, May 1995. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/336919/1stInquiryReport.pdf.
- Chambers, Simone. “Deliberative Democratic Theory.” *Annual Review of Political Science* 6, no. 1 (2003): 307–26. doi:10.1146/annurev.polisci.6.121901.085538.
- Charles, Cathy, and Suzanne DeMaio. “Lay Participation in Health Care Decision Making: A Conceptual Framework.” *Journal of Health Politics, Policy and Law* 18, no. 4 (1993): 881–904.
- Cheeseman, Graeme, and Hugh Smith. “Public Consultation or Political Choreography? The Howard Government’s Quest for Community Views on Defence Policy.” *Australian Journal of International Affairs* 55, no. 1 (2001): 83–100. doi:10.1080/10357710120055094.
- Civil Service. “Introduction to Sponsorship: An Induction Pack for New Sponsors of Arm’s-Length Bodies.” London: Civil Service, April 2014. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/303583/Introduction_to_Sponsorship_An_Induction_Handbook_for_New_Sponsors_of_Arm_s-Length_Bodies.pdf.
- Clarke, Simon. “Two Models of Ethics Committees.” *Journal of Bioethical Inquiry* 2, no. 1 (April 2005): 41–47. doi:10.1007/BF02448814.
- Clarke, Steve, and Rebecca Roache. “Introducing Transformative Technologies into Democratic Societies.” *Philosophy & Technology* 25, no. 1 (March 1, 2012): 27–45. doi:10.1007/s13347-011-0013-z.

- Clark, Stephen R. L. "Review of Religious Commitment and Secular Reason by Robert Audi." *Mind* 111, no. 443 (July 1, 2002): 639–43. doi:10.2307/3093625.
- Claxton, Karl, Mark Sculpher, Stephen Palmer, and Anthony J Culyer. "Causes for Concern: Is NICE Failing to Uphold Its Responsibilities to All NHS Patients?" *Health Economics* 24, no. 1 (January 1, 2015): 1–7. doi:10.1002/hec.3130.
- Claxton, K, S Martin, M Soares, N Rice, E Spackman, S Hinde, N Devlin, PC Smith, and M Sculpher. "Methods for the Estimation of the National Institute for Health and Care Excellence Cost-Effectiveness Threshold." *Health Technology Assessment* 19, no. 14 (February 18, 2015). doi:10.3310/hta19140.
- Coast, Joanna. "Is Economic Evaluation in Touch with Society's Health Values?" *BMJ* 329, no. 7476 (November 18, 2004): 1233–36. doi:10.1136/bmj.329.7476.1233.
- Cohen, Deborah. "Out of Joint: The Story of the ASR." *The BMJ* 342 (May 14, 2011): d2905. doi:10.1136/bmj.d2905.
- Cohen, Joshua. "Deliberation and Democratic Legitimacy." In *Deliberative Democracy: Essays on Reason and Politics*, edited by James Bohman and William Rehg. London: The MIT Press, 1997.
- . "Procedure and Substance in Deliberative Democracy." In *Deliberative Democracy: Essays on Reason and Politics*, edited by James Bohman and William Rehg, 407–37. London: The MIT Press, 1997.
- Commissioner for Public Appointments. "Code of Practice for Ministerial Appointments to Public Bodies." London, April 1, 2012. <http://publicappointmentscommissioner.independent.gov.uk/wp-content/uploads/2012/02/Code-of-Practice-20121.pdf>.
- Competition and Markets Authority. "Consumer Protection: Guidance to the CMA's Approach to Use of Its Consumer Powers," March 2014. <https://www.gov.uk/>

government/uploads/system/uploads/attachment_data/file/288624/CMA7_Consumer_Protection_guidance.pdf.

Counterpoint UK. “Sex Selection - Policy and Regulatory Review: A Report on the Key Findings from a Qualitative Research Study Prepared for the Human Fertilisation and Embryology Authority.” London, October 2002. [http://www.hfea.gov.uk/docs/Appendix_E_-_Qualitative_Research_Findings_report_of_research_conducted_by_Counterpoint_\(UK\)_Ltd.pdf](http://www.hfea.gov.uk/docs/Appendix_E_-_Qualitative_Research_Findings_report_of_research_conducted_by_Counterpoint_(UK)_Ltd.pdf).

Cowley, Christopher. “A New Rejection of Moral Expertise.” *Medicine, Health Care and Philosophy* 8, no. 3 (November 1, 2005): 273–79. doi:10.1007/s11019-005-1588-x.

———. “Expertise, Wisdom and Moral Philosophers: A Response to Gesang.” *Bioethics* 26, no. 6 (July 2012): 337–42. doi:10.1111/j.1467-8519.2010.01860.x.

Crosthwaite, Jan. “Moral Expertise: A Problem in the Professional Ethics of Professional Ethicists.” *Bioethics* 9, no. 5 (October 1995): 361–79.

Culyer, Anthony J. “NICE’s Use of Cost Effectiveness as an Exemplar of a Deliberative Process.” *Health Economics, Policy, and Law* 1, no. Pt 3 (July 2006): 299–318. doi:10.1017/S1744133106004026.

Dahl, Edgar. “Sex Selection: Laissez Faire or Family Balancing?” *Health Care Analysis: HCA: Journal of Health Philosophy and Policy* 13, no. 1 (March 2005): 87–90; discussion 91–93. doi:10.1007/s10728-005-2572-6.

Dakin, Helen Angela, Nancy J. Devlin, and Isaac A. O. Odeyemi. “‘Yes’, ‘No’ or ‘Yes, But’? Multinomial Modelling of NICE Decision-Making.” *Health Policy* 77, no. 3 (August 2006): 352–67. doi:10.1016/j.healthpol.2005.08.008.

Dakin, Helen, Nancy Devlin, Yan Feng, Nigel Rice, Phill O'Neill, and David Parkin.

“The Influence of Cost-Effectiveness and Other Factors on NICE Decisions.”

Health Economics, September 23, 2014. doi:10.1002/hec.3086.

Daniels, Norman. “Accountability for Reasonableness and the Citizens Council.” In *Patients, the Public and Priorities in Healthcare*, edited by Peter Littlejohns and Michael Rawlins. London: Radcliffe Publishing, 2009.

———. “Accountability for Reasonableness: Establishing a Fair Process for Priority Setting Is Easier than Agreeing on Principles.” *BMJ: British Medical Journal* 321, no. 7272 (November 25, 2000): 1300–1301.

———. “Normal Functioning and the Treatment-Enhancement Distinction.” *Cambridge Quarterly of Healthcare Ethics* 9, no. 03 (July 2000): 309–22. doi:null.

Daniels, Norman, and James E. Sabin. “Accountability for Reasonableness: An Update.” *BMJ* 337 (October 9, 2008): a1850. doi:10.1136/bmj.a1850.

———. *Setting Limits Fairly: Can We Learn to Share Medical Resources?* Oxford: Oxford University Press, 2002.

Davies, Celia, Margaret Wetherell, Elizabeth Barnett, and S. SEYMOUR-SMITH. “Opening the Box: Evaluating the Citizens Council of NICE: Report Prepared for the National Co-ordinating Centre for Research Methodology, NHS Research and Development Programme,” 2005. http://irep.ntu.ac.uk/R/?func=dbin-jump-full&object_id=188851&local_base=GEN01.

Deech, Ruth, and Anna Smajdor. *From IVF to Immortality: Controversy in the Era of Reproductive Technology*. Oxford: Oxford University Press, 2008.

Department for Business, Innovation and Skills. “Empowering and Protecting Consumers: Government Response to the Consultation on Institutional Reform,” April 2012. <https://www.gov.uk/government/uploads/system/uploads/>

attachment_data/file/253701/bis-12-510-empowering-protecting-consumers-government-response-1.pdf.

Department of Health. “Establishing New National Bodies - The Health and Social Care Act 2012 Fact Sheet,” June 15, 2012. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/138272/C7.-Establishing-new-national-bodies-270412.pdf.

———. “Government Response to the Consultation on Proposals to Transfer Functions from the Human Fertilisation and Embryology Authority and the Human Tissue Authority,” January 25, 2013. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/212742/Consultation_HFEA_and_HTA_government_response.pdf.

———. “Greater Accountability Locally and Nationally - The Health and Social Care Act 2012 Fact Sheet,” June 15, 2012. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/138264/B5.-Factsheet-greater-accountabilty-locally-and-nationally-250412.pdf.

———. “Liberating the NHS: Report of the Arm’s-Length Body Review,” July 26, 2010. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/216278/dh_118053.pdf.

———. “Modernising Health and Care Public Bodies - The Health and Social Care Act 2012 Fact Sheet,” June 15, 2012. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/138265/B6.-Factsheet-Streamlined-arms-length-bodies-250412.pdf.

———. “Review of the Human Fertilisation and Embryology Act: Proposals for Revised Legislation (including Establishment of the Regulatory Authority for

- Tissue and Embryos)." Cm 6989, December 2006. http://www.hfea.gov.uk/docs/Review_HFEA_Act_White_Paper_DH.pdf.
- . "The Health and Social Care Act 2012 Overview Fact Sheet," June 15, 2012. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/138257/A1.-Factsheet-Overview-240412.pdf.
- . "The Human Fertilisation and Embryology Act 1990: As Amended - an Illustrative Text," November 9, 2007. http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalasset/s/@dh/@en/documents/digitalasset/dh_080206.pdf.
- Deutscher Ethikrat. "About Us: Ethics Council Act," n.d. <http://www.ethikrat.org/about-us/ethics-council-act>.
- Dixon-Woods, Mary, Shona Agarwal, David Jones, Bridget Young, and Alex Sutton. "Synthesising Qualitative and Quantitative Evidence: A Review of Possible Methods." *Journal of Health Services Research & Policy* 10, no. 1 (January 1, 2005): 45–53B.
- Dodds, Susan, and Colin Thomson. "Bioethics and Democracy: Competing Roles of National Bioethics Organisations." *Bioethics* 20, no. 6 (2006): 326–38.
- Dommett, Katharine. "Triennial Reviews and Public Body Reform in the UK." ESRC Shrinking the State Project, January 2014.
- Duerden, Martin. "From a Cancer Drug Fund to Value Based Pricing of Drugs." *BMJ* 341 (August 12, 2010): c4388. doi:10.1136/bmj.c4388.
- Duncan, Peter. "Health, Health Care and the Problem of Intrinsic Value." *Journal of Evaluation in Clinical Practice* 16, no. 2 (April 2010): 318–22. doi:10.1111/j.1365-2753.2010.01392.x.

- Dyer, Clare. "Independent Review Is Needed to Decide HFEA's Fate." *BMJ* 342 (January 19, 2011): d379. doi:10.1136/bmj.d379.
- . "New Life for the HFEA?" *BMJ* 342 (January 25, 2011): d439. doi:10.1136/bmj.d439.
- Edwards, Kyle T. "Methods of Legitimation: How Ethics Committees Decide Which Reasons Count in Public Policy Decision-Making." *Social Science & Medicine* (1982) 113 (July 2014): 34–41. doi:10.1016/j.socscimed.2014.04.043.
- Edwards, Kyle T., and Zuzanna Deans. "Empirical Ethics and the Role of the Professional Ethicist in Public Policymaking." In *Empirical Bioethics: Practical and Theoretical Perspectives*, edited by Jon Ives, Michael Dunn, and Alan Cribb. Expected. Cambridge: Cambridge University Press, expected 2015.
- Edwards, Kyle T., and Michael Dunn. "Ethics Expertise in Public Policymaking." *Journal of Medicine and Philosophy*, forthcoming Special Issue on Ethics Expertise (2016).
- Edwards, Kyle T., and Caroline J. Huang. "Bridging the Consumer-Medical Divide: How to Regulate Direct-to-Consumer Genetic Testing." *Hastings Center Report* 44, no. 3 (May 1, 2014): 17–19. doi:10.1002/hast.310.
- Elster, Jon. "The Market and the Forum: Three Varieties of Political Theory." In *Deliberative Democracy: Essays on Reason and Politics*, edited by James Bohman and William Rehg, 3–33. London: The MIT Press, 1997.
- Encyclopædia Britannica Online*. s. v. "Board of Trade." Accessed February 19, 2014. <http://www.britannica.com/EBchecked/topic/601629/Board-of-Trade>.
- Entwistle, Vikki A., Mary J. Renfrew, Steven Yearley, John Forrester, and Tara Lamont. "Lay Perspectives: Advantages for Health Research." *BMJ* 316, no. 7129 (February 7, 1998): 463–66. doi:10.1136/bmj.316.7129.463.

- Epstein, Steven. "The Construction of Lay Expertise: AIDS Activism and the Forging of Credibility in the Reform of Clinical Trials." *Science, Technology & Human Values* 20, no. 4 (October 1, 1995): 408–37. doi:10.1177/016224399502000402.
- Estlund, David M. *Democratic Authority: A Philosophical Framework*. Princeton, N.J.: Princeton University Press, 2009.
- . "Who's Afraid of Deliberative Democracy - On the Strategic/Deliberative Dichotomy in Recent Constitutional Jurisprudence." *Texas Law Review* 71 (1993 1992): 1437.
- European Commission. "Executive Summary of the Impact Assessment on the Revision of the Regulatory Framework for Medical Devices." Commission Staff Working Document. Brussels, September 26, 2012. http://ec.europa.eu/growth/sectors/medical-devices_old/documents/revision/files/revision_docs/revision_ia_summary_en.pdf.
- European Commission. "Commission Staff Working Document: Impact Assessment on the Revision of the Regulatory Framework for Medical Devices." Brussels, 2012. <http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52012SC0273&from=EN>.
- . "Council Directive 93/42/EEC Concerning Medical Devices," June 14, 1993. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:1993L0042:20071011:EN:PDF%20>.
- . "Directive 2001/83/EC of the European Parliament and of the Council: On the Community Code Relating to Medicinal Products for Human Use," November 6, 2011. http://ec.europa.eu/health/files/eudralex/vol-1/dir_2001_83_consol_2012/dir_2001_83_consol_2012_en.pdf.

- . “Exploratory Process on the Future of the Medical Devices Sector: Potential Themes for Further Reflection at the European Level and Issues Identified by the Members.” Brussels, January 28, 2010. http://ec.europa.eu/health/medical-devices/files/exploratory_process/final_report_en.pdf.
- . “Guidance Document on the Demarcation Between the Cosmetic Products Directive 76/768 and the Medicinal Products Directive 2001/83 as Agreed Between the Commission Services and the Competent Authorities of Member States.” Brussels, April 30, 2004. http://ec.europa.eu/growth/sectors/cosmetics/products/borderline-products/docs/guidance_doc_cosm-medicinal_en.pdf.
- . “Medical Devices: Guidance Document - Classification of Medical Devices.” Guidance relating to the application of the Council Directive 93/42/EEC on medical devices. Brussels, June 2010. http://ec.europa.eu/health/medical-devices/files/meddev/2_4_1_rev_9_classification_en.pdf.
- . “Proposal for a Regulation of the European Parliament and of the Council: On Medical Devices, and Amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009.” Brussels, September 9, 2012. http://ec.europa.eu/growth/sectors/medical-devices_old/documents/revision/files/revision_docs/proposal_2012_542_en.pdf.
- . “Revisions of Medical Device Directives,” June 10, 2014. http://ec.europa.eu/growth/sectors/medical-devices/regulatory-framework/revision/index_en.htm.
- . “Safe, Effective and Innovative Medical Devices and in Vitro Diagnostic Medical Devices for the Benefit of Patients, Consumers and Healthcare Professionals.” Communication from the Commission to the European

- Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. Brussels, September 26, 2012. <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52012DC0540&from=EN>.
- Evans, David. "Hierarchy of Evidence: A Framework for Ranking Evidence Evaluating Healthcare Interventions." *Journal of Clinical Nursing* 12, no. 1 (January 2003): 77–84.
- Evans, Robert, and Alexandra Plows. "Listening Without Prejudice? Re-Discovering the Value of the Disinterested Citizen." *Social Studies of Science* 37, no. 6 (December 1, 2007): 827–53. doi:10.1177/0306312707076602.
- "Evidence." *Oxford Dictionaries*. Oxford University Press, August 14, 2015. <http://www.oxforddictionaries.com/definition/english/evidence>.
- Facey, Karen, Antoine Boivin, Javier Gracia, Helle Ploug Hansen, Alessandra Lo Scalzo, Jean Mossman, and Ann Single. "Patients' Perspectives in Health Technology Assessment: A Route to Robust Evidence and Fair Deliberation." *International Journal of Technology Assessment in Health Care* 26, no. 03 (July 2010): 334–40. doi:10.1017/S0266462310000395.
- Fefer, Enrique. "Chapter 6: Pharmaceutical Legislation and Regulation." In *MDS-3: Managing Access to Medicines and Health Technologies*. Arlington: Management Sciences for Health, 2012. <https://www.msh.org/sites/msh.org/files/mds3-jan2014.pdf>.
- Feinstein, Alvan R, and Ralph I Horwitz. "Problems in the 'Evidence' of 'Evidence-Based Medicine.'" *The American Journal of Medicine* 103, no. 6 (December 1997): 529–35. doi:10.1016/S0002-9343(97)00244-1.

- Ferejohn, John. "Incumbent Performance and Electoral Control." *Public Choice* 50, no. 1–3 (January 1986): 5–25. doi:10.1007/BF00124924.
- "Fertility Watchdog 'Incompetent.'" *BBC*, December 10, 2004, sec. Health. <http://news.bbc.co.uk/1/hi/health/4084365.stm>.
- Fish, Stanley. "Mutual Respect as a Device of Exclusion." In *Deliberative Politics: Essays on Democracy and Disagreement*, edited by Stephen Macedo. Oxford: Oxford University Press, 1999.
- Flinders, Matthew. *Delegated Governance and the British State: Walking without Order*. Oxford: Oxford University Press, 2008.
- Flinders, Matthew, Katharine Dommett, and Katherine Tonkiss. "Bonfires and Barbecues: Coalition Governance and the Politics of Quango Reform." *Contemporary British History* 28, no. 1 (January 2, 2014): 56–80. doi:10.1080/13619462.2014.882776.
- Flinders, Matthew, and Chris Skelcher. "Shrinking the Quango State: Five Challenges in Reforming Quangos." *Public Money & Management* 32, no. 5 (2012): 327–34. doi:10.1080/09540962.2012.703410.
- Flöel, Agnes. "tDCS-Enhanced Motor and Cognitive Function in Neurological Diseases." *Neuro-Enhancement* 85, Part 3 (January 15, 2014): 934–47. doi:10.1016/j.neuroimage.2013.05.098.
- Florin, Dominique, and Jennifer Dixon. "Public Involvement in Health Care." *BMJ: British Medical Journal* 328, no. 7432 (January 17, 2004): 159–61.
- Forster, Rudolf, and Jonathan Gabe. "Voice or Choice? Patient and Public Involvement in the National Health Service in England under New Labour." *International Journal of Health Services* 38, no. 2 (January 1, 2008): 333–56. doi:10.2190/HS.38.2.g.

- “Framework Agreement between the Department of Health and the Human Fertilisation and Embryology Authority,” 2014. [http://www.hfea.gov.uk/docs/HFEA_Framework_Agreement_v16_\(FINAL_signed_for_publication\).pdf](http://www.hfea.gov.uk/docs/HFEA_Framework_Agreement_v16_(FINAL_signed_for_publication).pdf).
- “Framework Agreement between the Department of Health and the National Institute for Health and Care Excellence,” 2014. <https://www.nice.org.uk/Media/Default/About/Who-we-are/NICE-and-Dept-of-Health-Framework-Agreement-March-2014.pdf>.
- Freeman, Samuel. “Deliberative Democracy: A Sympathetic Comment.” *Philosophy & Public Affairs* 29, no. 4 (October 1, 2000): 371–418. doi:10.1111/j.1088-4963.2000.00371.x.
- Fregni, F., M. A. Nitsche, C. K. Loo, A. R. Brunoni, P. Marangolo, J. Leite, S. Carvalho, et al. “Regulatory Considerations for the Clinical and Research Use of Transcranial Direct Current Stimulation (tDCS): Review and Recommendations from an Expert Panel.” *Clinical Research and Regulatory Affairs* 32, no. 1 (December 2, 2014): 22–35. doi:10.3109/10601333.2015.980944.
- Friedman, Alex. “Beyond Accountability for Reasonableness.” *Bioethics* 22, no. 2 (February 1, 2008): 101–12. doi:10.1111/j.1467-8519.2007.00605.x.
- Galston, William. “Diversity, Toleration, and Deliberative Democracy: Religious Minorities and Public Schooling.” In *Deliberative Politics: Essays on Democracy and Disagreement*, edited by Stephen Macedo. Oxford: Oxford University Press, 1999.
- Ganyo, Michelle, Michael Dunn, and Tony Hope. “Ethical Issues in the Use of Fall Detectors.” *Ageing & Society* 31, no. 08 (November 2011): 1350–67. doi:10.1017/S0144686X10001443.

- Gash, Tom, Sir Ian Magee, Jill Rutter, and Nicole Smith. "Read before Burning: Arm's Length Government for a New Administration." Institute for Government, July 2010.
- Gash, Tom, and Jill Rutter. "The Quango Conundrum." *The Political Quarterly* 82, no. 1 (2011): 95–101. doi:10.1111/j.1467-923X.2011.02167.x.
- Gaus, Gerald. *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World*. Cambridge University Press, 2010.
- Gaus, Gerald F. *Justificatory Liberalism: An Essay on Epistemology and Political Theory*. New York: Oxford University Press, 1996.
- Gentleman, Amelia. "A Quiet Battle for Life: IVF Regulator Aims to Show Its Worth." *The Guardian*, January 10, 2011, sec. Politics. <http://www.theguardian.com/politics/2011/jan/10/battle-life-ivf-regulator-hfea>.
- George, Robert P. "Law, Democracy, and Moral Disagreement." *Harvard Law Review* 110, no. 7 (May 1, 1997): 1388–1406. doi:10.2307/1342176.
- . "Public Reason and Political Conflict: Abortion and Homosexuality." *The Yale Law Journal* 106, no. 8 (June 1, 1997): 2475–2504. doi:10.2307/797225.
- Gesang, Bernward. "Are Moral Philosophers Moral Experts?" *Bioethics* 24, no. 4 (May 2010): 153–59.
- Goldenberg, Maya J. "On Evidence and Evidence-Based Medicine: Lessons from the Philosophy of Science." *Part Special Issue: Gift Horse or Trojan Horse? Social Science Perspectives on Evidence-Based Health Care* *Part Special Issue: Gift Horse or Trojan Horse? Social Science Perspectives on Evidence-Based Health Care* 62, no. 11 (June 2006): 2621–32. doi:10.1016/j.socscimed.2005.11.031.

- Goodin, Robert E., and John S. Dryzek. "Deliberative Impacts: The Macro-Political Uptake of Mini-Publics." *Politics & Society* 34, no. 2 (June 1, 2006): 219–44. doi:10.1177/0032329206288152.
- "Government Opens the Doors on Quangos." *Local Government Chronicle*, June 30, 1998. <http://www.lgcplus.com/government-opens-the-doors-on-quangos/1457666.article>.
- Greely, Hank. "Of Science, CRISPR-Cas9, and Asilomar." *Law and Biosciences Blog*, April 4, 2015. <https://blogs.law.stanford.edu/lawandbiosciences/2015/04/04/of-science-crispr-cas9-and-asilomar/>.
- Greenawalt, Kent. *Private Consciences and Public Reasons*. New York: Oxford University Press, 1995.
- Green, Ronald M. "For Richer or Poorer? Evaluating the President's Council on Bioethics." *HEC Forum* 18, no. 2 (August 8, 2006): 108–24. doi:10.1007/s10730-006-9001-6.
- Greve, Carsten, Matthew Flinders, and Sandra Van Thiel. "Quangos—What's in a Name? Defining Quangos from a Comparative Perspective." *Governance* 12, no. 2 (1999): 129–46. doi:10.1111/0952-1895.951999095.
- Gutmann, Amy, and Dennis Thompson. *Democracy and Disagreement*. Cambridge: The Belknap Press of Harvard University Press, 1996.
- Gutmann, Amy, and Dennis Thomson. *Why Deliberative Democracy?* Princeton: Princeton University Press, 2004.
- Habermas, Jurgen. "Reconciliation Through the Public Use of Reason: Remarks on John Rawls's Political Liberalism." *The Journal of Philosophy* 92, no. 3 (March 1, 1995): 109–31. doi:10.2307/2940842.

- Habermas, Jürgen. *The Theory of Communicative Action: Reason and the Rationalization of Society*. Translated by Thomas McCarthy. Boston: Beacon Press, 1985.
- Hahn, Jim, and Christopher M. Davis. “The Independent Payment Advisory Board.” CRS Report for Congress. Washington, DC: Congressional Research Service, April 17, 2013.
- Harris, J. “Double Jeopardy and the Veil of Ignorance--a Reply.” *Journal of Medical Ethics* 21, no. 3 (June 1, 1995): 151–57. doi:10.1136/jme.21.3.151.
- . “QALYfying the Value of Life.” *Journal of Medical Ethics* 13, no. 3 (September 1, 1987): 117–23. doi:10.1136/jme.13.3.117.
- Harris, John. “It’s Not NICE to Discriminate.” *Journal of Medical Ethics* 31, no. 7 (July 1, 2005): 373–75. doi:10.1136/jme.2005.012906.
- . “Nice and Not so Nice.” *Journal of Medical Ethics* 31, no. 12 (December 1, 2005): 685–88. doi:10.1136/jme.2005.014134.
- . “Sex Selection and Regulated Hatred.” *Journal of Medical Ethics* 31, no. 5 (May 1, 2005): 291–94. doi:10.1136/jme.2003.007526.
- HC Deb 14 October 2010 vol 516 col 26WS – col 28WS.
- Health and Social Care Act 2012*, c. 7. <http://www.legislation.gov.uk/ukpga/2012/7/contents>.
- Health Select Committee. *National Institute for Clinical Excellence*. 16 January 2013. HC 782-I 2012-2013.
- Heneghan, C. “The Saga of Poly Implant Prosthese Breast Implants.” *BMJ* 344, no. jan11 2 (January 11, 2012): e306–e306. doi:10.1136/bmj.e306.
- Herissone-Kelly, Peter. “The Prohibition of Sex Selection for Social Reasons in the United Kingdom: Public Opinion Trumps Reproductive Liberty?” *Cambridge*

- Quarterly of Healthcare Ethics* 15, no. 03 (July 2006): 261–72.
doi:10.1017/S0963180106060336.
- HFEA Ethics and Law Committee. “Ethical Values and the HFEA,” December 2003.
http://www.hfea.gov.uk/docs/ELC_2_Dec03.pdf.
- . “Ethics Framework,” February 2011. [http://www.hfea.gov.uk/docs/2011-06-15_-_ELAC_ethics_framework_for_website_\(2\).pdf](http://www.hfea.gov.uk/docs/2011-06-15_-_ELAC_ethics_framework_for_website_(2).pdf).
- . “HFEA Principles.” London: HFEA, July 2003.
http://www.hfea.gov.uk/docs/ELC_1_july03.pdf.
- HFEA Ethics and Standards Committee. “Minutes of the Proceedings of the HFEA Ethics and Standards Committee,” January 14, 2015.
http://www.hfea.gov.uk/docs/Ethics_and_Standards_Committee_Meeting_-_Signed_Minutes_-_Jan_15.pdf.
- Hill, Suzanne, Silvio Garattini, Jos van Loenhout, Bernie J. O’Brien, and Kees de Joncheere. “Technology Appraisal Programme of the National Institute for Clinical Excellence: A Review by WHO.” Copenhagen: WHO Regional Office for Europe, 2003. <http://apps.who.int/medicinedocs/documents/s19194en/s19194en.pdf>.
- Hogg, Christine, and Charlotte Williamson. “Whose Interests Do Lay People Represent? Towards an Understanding of the Role of Lay People as Members of Committees.” *Health Expectations: An International Journal of Public Participation in Health Care and Health Policy* 4, no. 1 (March 2001): 2–9.
- Holm, Soren. “The Second Phase of Priority Setting. Goodbye to the Simple Solutions: The Second Phase of Priority Setting in Health Care.” *BMJ (Clinical Research Ed.)* 317, no. 7164 (October 10, 1998): 1000–1002.

- Hughes, David, Caroline Mullen, and Peter Vincent-Jones. "Choice vs. Voice? PPI Policies and the Re-Positioning of the State in England and Wales." *Health Expectations: An International Journal of Public Participation in Health Care and Health Policy* 12, no. 3 (September 2009): 237–50. doi:10.1111/j.1369-7625.2009.00559.x.
- Hughes, J. C. "Electronic Tagging of People with Dementia Who Wander." *BMJ* 325, no. 7369 (October 19, 2002): 847–48. doi:10.1136/bmj.325.7369.847.
- Hughes, Julian C, and Stephen J Louw. "Electronic Tagging of People with Dementia Who Wander." *BMJ: British Medical Journal* 325, no. 7369 (October 19, 2002): 847–48.
- Human Fertilisation and Embryology Authority. "Annual Report and Accounts 2009/10," July 14, 2010. http://www.hfea.gov.uk/docs/Annual_Report_2010_Web_reduced.pdf.
- . "Choices and Boundaries Report: A Summary of Responses to the HFEA Public Discussion," 2006. http://www.hfea.gov.uk/docs/Choices_and_boundaries_Report_2006_summary.pdf.
- . "Decision Tree for Use by Licence Committee and Executive Licensing Panel Consideration of Applications to Grant or New a Licence," n.d. http://www.hfea.gov.uk/docs/2014_004645__2014-03-06_-_decision_trees_collection_-_A3_size_pdf_-_final.pdf.
- . "Ethics and Law Advisory Committee (ELAC)," December 5, 2013. <http://www.hfea.gov.uk/ELAC.html>.
- . "Mitochondria Replacement Consultation: Advice to Government," March 2013. http://www.hfea.gov.uk/docs/Mitochondria_replacement_consultation_-_advice_for_Government.pdf.

- . “Our Role as Regulator,” April 8, 2015. <http://www.hfea.gov.uk/135.html>.
- . “Pre-Implantation - Information on ‘Saviour Siblings’ and Testing for Genetic Conditions,” April 1, 2014. <http://www.hfea.gov.uk/preimplantation-tissue-typing.html>.
- . “Sex Selection: Choice and Responsibility in Human Reproduction,” January 2003.
http://www.hfea.gov.uk/docs/Sex_Selection_choice_and_responsibility.pdf.
- . “Sex Selection: Options for Regulation - A Report on the HFEA’s 2002-03 Review of Sex Selection Including a Discussion of Legislative and Regulatory Options,” 2003. http://www.hfea.gov.uk/docs/Final_sex_selection_main_report.pdf.
- . “Sex Selection: Public Consultation Document,” January 1993.
http://www.hfea.gov.uk/docs/Sex_Selection_Consultation_1993.pdf.
- Ipsos MORI. “Risks and Benefits of Medicines and Medical Devices – Perceptions, Communication and Regulation: General Public Quantitative Survey.” Report for the MHRA, 2009. <http://www.mhra.gov.uk/home/groups/comms-ic/documents/websiteresources/con052027.pdf>.
- Ives, Jonathan, Sarah Damery, and Sabi Redwod. “PPI, Paradoxes and Plato: Who’s Sailing the Ship?” *Journal of Medical Ethics* 39, no. 3 (March 1, 2013): 181–85.
doi:10.1136/medethics-2011-100150.
- Jackson, Emily. *Law and the Regulation of Medicines*. Oxford: Hart Publishing Ltd, 2012.
- Jeremy Howick, Iain Chalmers, Paul Glasziou, Trish Greenhalgh, Carl Heneghan, Alessandro Liberati, Ivan Moschetti, Bob Phillips, and Hazel Thornton. “The 2011 Oxford CEBM Levels of Evidence (Introductory Document).” Oxford

Centre for Evidence-Based Medicine, n.d.

<http://www.cebm.net/index.aspx?o=5653>.

Kadosh, Roi Cohen. “Using Transcranial Electrical Stimulation to Enhance Cognitive Functions in the Typical and Atypical Brain.” *Translational Neuroscience* 4, no. 1 (March 7, 2013): 20–33. doi:10.2478/s13380-013-0104-7.

Kamm, Frances M. “The Philosopher as Insider and Outsider.” *Journal of Medicine and Philosophy* 15, no. 4 (August 1, 1990): 347–74. doi:10.1093/jmp/15.4.347.

Kelly, Erin, and Lionel McPherson. “On Tolerating the Unreasonable.” *Journal of Political Philosophy* 9, no. 1 (March 1, 2001): 38–55. doi:10.1111/1467-9760.00117.

Kennedy, Ian. “Patients Are Experts in Their Own Field.” *BMJ: British Medical Journal* 326, no. 7402 (June 14, 2003): 1276–77.

Kent Greenawalt. *Religious Convictions and Political Choice*. New York: Oxford University Press, 1991.

Kieslich, Katharina, Peter Littlejohns, Albert Weale, and Sarah Clark. “Does Accountability for Reasonableness Work? A Protocol for a Mixed Methods Study Using an Audit Tool to Evaluate the Decision-Making of Clinical Commissioning Groups in England.” *BMJ Open* 5, no. 7 (July 1, 2015): e007908. doi:10.1136/bmjopen-2015-007908.

Knechtel, Lilly, Renate Thienel, and Ulrich Schall. “Transcranial Direct Current Stimulation: Neurophysiology and Clinical Applications.” *Neuropsychiatry* 3, no. 1 (February 1, 2013): 89–96. doi:10.2217/npv.12.78.

Koerner, Brendan. “Embryo Police.” *Wired Magazine*, February 2002. <http://archive.wired.com/wired/archive/10.02/biotech.html>.

- Kuersten, Andreas, and Roy H. Hamilton. "The Brain, Cognitive Enhancement Devices, and European Regulation." *Journal of Law and the Biosciences*, August 26, 2014, lsu019. doi:10.1093/jlb/lsu019.
- Lachenmeier, Dirk W., Christian Steffen, Oliver el-Atma, Sibylle Maixner, Sigrid Löbell-Behrends, and Matthias Kohl-Himmelseher. "What Is a Food and What Is a Medicinal Product in the European Union? Use of the Benchmark Dose (BMD) Methodology to Define a Threshold for 'pharmacological Action.'" *Regulatory Toxicology and Pharmacology* 64, no. 2 (November 2012): 286–95. doi:10.1016/j.yrtph.2012.08.017.
- Lehoux, P., G. Daudelin, and J. Abelson. "The Unbearable Lightness of Citizens within Public Deliberation Processes." *Social Science & Medicine* 74, no. 12 (June 2012): 1843–50. doi:10.1016/j.socscimed.2012.02.023.
- Lenaghan, Jo. "Involving the Public in Rationing Decisions. The Experience of Citizens Juries." *Health Policy* 49, no. 1–2 (September 1999): 45–61. doi:10.1016/S0168-8510(99)00042-1.
- Lever, Annabelle. "Democracy and Lay Participation: The Case of NICE." In *The Future of Public Service Reform*, edited by Henry Kippin Gerry Stoker. bloomsbury academic press, 2013.
- Levy, N. "Deafness, Culture, and Choice." *Journal of Medical Ethics* 28, no. 5 (October 1, 2002): 284–85. doi:10.1136/jme.28.5.284.
- Majone, Giandomenico. "Europe's 'Democratic Deficit': The Question of Standards." *European Law Journal* 4, no. 1 (March 1, 1998): 5–28. doi:10.1111/1468-0386.00040.
- . "The Rise of the Regulatory State in Europe." *West European Politics* 17, no. 3 (July 1, 1994): 77–101. doi:10.1080/01402389408425031.

- Mansbridge, Jane. "Using Power/Fighting Power: The Polity." In *Democracy and Difference: Contesting the Boundaries of the Political*, edited by Seyla Benhabib. Princeton, N.J: Princeton University Press, 1996.
- Mary Ann Elston. "The Politics of Professional Power: Medicine in a Changing Health Service." In *The Sociology of the Health Service*, edited by Michael Bury, Michael Calnan, and Jonathan Gabe. London: Routledge, 2002.
- Maskin, Eric, and Jean Tirole. "The Politician and the Judge: Accountability in Government." *American Economic Review* 94, no. 4 (2004): 1034–54.
- Maslen, Hannah, Thomas Douglas, Roi Cohen Kadosh, Neil Levy, and Julian Savulescu. "The Regulation of Cognitive Enhancement Devices: Extending the Medical Model." *Journal of Law and the Biosciences* 1, no. 1 (March 1, 2014): 68–93. doi:10.1093/jlb/lst003.
- Mason, Helen, Michael Jones-Lee, and Cam Donaldson. "Modelling the Monetary Value of a QALY: A New Approach Based on UK Data." *Health Economics* 18, no. 8 (August 1, 2009): 933–50. doi:10.1002/hec.1416.
- Mays, Nicholas, Catherine Pope, and Jennie Popay. "Systematically Reviewing Qualitative and Quantitative Evidence to Inform Management and Policy-Making in the Health Field." *Journal of Health Services Research & Policy* 10, no. suppl 1 (July 15, 2005): 6–20. doi:10.1258/1355819054308576.
- McCabe, Christopher, Karl Claxton, and Anthony J. Culyer. "The NICE Cost-Effectiveness Threshold: What It Is and What That Means." *Pharmacoeconomics* 26, no. 9 (2008): 733–44.
- McCracken, Justin. "Review of the Human Fertilisation and Embryology Authority and the Human Tissue Authority: An Independent Report to the Parliamentary Under Secretary of State for Public Health and the Minister for the Cabinet Office,"

April 2013. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/216947/Justin_McCracken_report_of_review_of_HFEA_and_HTA.pdf.

Medicines and Healthcare Products Regulatory Agency. “A Guide to What Is a Medicinal Product.” MHRA Guidance Note No. 8, November 2012. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/398998/A_guide_to_what_is_a_medicinal_product.pdf.

———. “Borderlines between Medical Devices and Medicinal Products.” Guidance on legislation, June 2013. <https://www.gov.uk/government/publications/borderlines-between-medical-devices-and-medicinal-products>.

———. “GVK Biosciences: MHRA Responds to European Medicines Agency’s Recommendation to Suspend Medicines.” *Press Release*, January 23, 2015. <https://www.gov.uk/government/news/gvk-biosciences-mhra-responds-to-european-medicines-agencys-recommendation-to-suspend-medicines>.

———. “Medical Device Stand-Alone Software Including Apps.” Guidance, August 8, 2014. <https://www.gov.uk/government/publications/medical-devices-software-applications-apps/medical-device-stand-alone-software-including-apps#more-information>.

———. “Medicines and Medical Devices Regulation: What You Need to Know.” London: MHRA, April 2008. <http://www.mhra.gov.uk/home/groups/comms-ic/documents/websiteresources/con2031677.pdf>.

———. “Metal-on-Metal (MoM) Hip Replacements - Updated Advice with Patient Follow Ups.” *Medical Safety Alert*, June 25, 2012. <https://www.gov.uk/drug-device-alerts/medical-device-alert-metal-on-metal-mom-hip-replacements-updated-advice-with-patient-follow-ups>.

- . “MHRA Response to the Independent Review on Access to Clinical Advice and Engagement With the Clinical Community in Relation to Medical Devices,” July 18, 2014. <http://webarchive.nationalarchives.gov.uk/20141205150130/http://www.mhra.gov.uk/home/groups/comms-po/documents/news/con437436.pdf>.
- . “New Chair Appointed for the Newly-Formed Devices Expert Advisory Committee.” Press Release, February 18, 2015. <https://www.gov.uk/government/news/new-chair-appointed-for-the-newly-formed-devices-expert-advisory-committee>.
- . “New Independent Expert Advisory Group to Be Created as the MHRA Responds to Stephenson Review.” Press Release, July 18, 2014. <http://webarchive.nationalarchives.gov.uk/20141205150130/http://www.mhra.gov.uk/NewsCentre/Pressreleases/CON437435>.
- . “Patient and Public Engagement EAG.” *Membership - Commission on Human Medicines*. Accessed August 19, 2015. <https://www.gov.uk/government/organisations/commission-on-human-medicines/about/membership#patient-and-public-engagement-eag>.
- . “The Yellow Card Scheme: Guidance for Healthcare Professionals,” January 23, 2015. <https://www.gov.uk/the-yellow-card-scheme-guidance-for-healthcare-professionals>.
- . “What Principles Inform the Decisions?,” February 7, 2008. <http://webarchive.nationalarchives.gov.uk/20080804194642/mhra.gov.uk/printpreview/defaultsplashpp/con2032522?dynamiclistquery=null&dynamiclistsortby=null&dynamiclistsortorder=null&dynamiclisttitle=null&pagenumber=1&title=what%20principles%20inform%20the%20decisions%3F&resultcount=10>.

- . “What Questions Should Be Considered to Ensure That Reasonable Decisions Are Made?,” February 7, 2008. <http://webarchive.nationalarchives.gov.uk/20080804194642/http://mhra.gov.uk/Howweregulate/Whatquestionsshouldbeconsideredtoensurethatreasonabledecisionsaremade/index.htm>.
- . “What We Do.” *About Us*. Accessed August 16, 2015. <https://www.gov.uk/government/organisations/medicines-and-healthcare-products-regulatory-agency/about>.
- Melnitchouk, Olya, and Alison McAdams. “New Regulations Simplify 40 Years of Medicines Law.” *Health Service Journal*, November 29, 2012. <http://www.hsj.co.uk/resource-centre/best-practice/qipp-resources/new-regulations-simplify-40-years-of-medicines-law/5051815.article>.
- “Memorandum of Understanding (MOU) Between National Institute for Health and Care Excellence (NICE) and Patients Involved in NICE (PIN),” November 2013. <https://www.nice.org.uk/media/default/About/NICE-Communities/Public-involvement/PIN/NICE-PIN-memorandum-of-understanding-November-2013.pdf>.
- MHRA Committee on the Safety of Devices. “Minutes of the Committee on the Safety of Devices,” March 20, 2014. http://www.gov.uk/government/uploads/system/uploads/attachment_data/file/401677/20-03-14_-_CSD_final_approved.pdf.
- Michelman, Frank I. “How Can the People Ever Make the Law? A Critique of Deliberative Democracy.” In *Deliberative Democracy: Essays on Reason and Politics*, edited by James Bohman and William Rehg. London: The MIT Press, 1997.

- Milewa, Timothy. "Health Technology Adoption and the Politics of Governance in the UK." *Social Science & Medicine* 63, no. 12 (December 2006): 3102–12. doi:10.1016/j.socscimed.2006.08.009.
- . "Representation and Legitimacy in Health Policy Formulation at a National Level: Perspectives from a Study of Health Technology Eligibility Procedures in the United Kingdom." *Health Policy (Amsterdam, Netherlands)* 85, no. 3 (March 2008): 356–62. doi:10.1016/j.healthpol.2007.09.001.
- Milewa, Timothy, and Christine Barry. "Health Policy and the Politics of Evidence." *Social Policy & Administration* 39, no. 5 (October 1, 2005): 498–512. doi:10.1111/j.1467-9515.2005.00452.x.
- Mill, John Stuart. *On Liberty*. Edited by Steven M. Cahn. Rowman & Littlefield, 2005.
- Mitton, Craig R., Meghan McMahon, Steve Morgan, and Jennifer Gibson. "Centralized Drug Review Processes: Are They Fair?" *Social Science & Medicine (1982)* 63, no. 1 (July 2006): 200–211. doi:10.1016/j.socscimed.2005.11.049.
- Montgomery, Jonathan. "Public Ethics and Faith." *Theology* 117, no. 5 (September 1, 2014): 342–48. doi:10.1177/0040571X14537031.
- Moore, Alfred. "Public Bioethics and Deliberative Democracy." *Political Studies* 58, no. 4 (2010): 715–30. doi:10.1111/j.1467-9248.2010.00836.x.
- . "Public Bioethics and Public Engagement: The Politics of 'Proper Talk.'" *Public Understanding of Science (Bristol, England)* 19, no. 2 (March 2010): 197–211.
- Mullen, Caroline. "Representation or Reason: Consulting the Public on the Ethics of Health Policy." *Health Care Analysis* 16, no. 4 (December 1, 2008): 397–409. doi:10.1007/s10728-007-0068-2.

- National Collaboration for Integrated Care and Support. “Integrated Care and Support: Our Shared Commitment,” May 2013. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/287815/DEFINITIVE_FINAL_VERSION_Integrated_Care_and_Support_-_Our_Shared_Commitment_2013-05-13.pdf.
- National Health and Medical Research Council. “About the NHMRC,” May 21, 2015. <https://www.nhmrc.gov.au/about>.
- . “An Invitation to Make a Submission: Review of Part B of the Ethical Guidelines on the Use of Assisted Reproductive Technology in Clinical Practice and Research,” July 15, 2015. <http://consultations.nhmrc.gov.au/files/consultations/drafts/artdraftethicalguidelines150722.pdf>.
- . “Australian Health Ethics Committee (AHEC),” July 1, 2015. <https://www.nhmrc.gov.au/about/nhmrc-committees/australian-health-ethics-committee-ahec>.
- . “Review of Human Cloning and Human Embryo Research Legislation (2005),” July 17, 2015. <https://www.nhmrc.gov.au/about/nhmrc-committees/embryo-research-licensing-committee/human-embryos-and-cloning/review-human-cl>.
- National Institute for Health and Care Excellence. “Appraising Life-Extending, End of Life Treatments.” Supplementary Advice to Appraisal Committees, July 2009.
- . “Comments from Novartis Pharmaceuticals UK Limited on the Appraisal Consultation Document (ACD) for the Health Technology Appraisal of Everolimus for the Second-Line Treatment of Metastatic Renal Cell Carcinoma,” March 2010. <https://www.nice.org.uk/guidance/ta219/documents/manufacturers-response-to-the-acd2>.
- . “Decision of the Appeal Panel: Advice on Cetuximab for the Treatment of Locally Advanced Squamous Cell Cancer of the Head and Neck,” August 22,

2007. <https://www.nice.org.uk/guidance/ta145/documents/head-and-neck-cancer-cetuximab-result-of-appeal-2>.
- . “Decision of the Appeal Panel: Advice on Everolimus for the Second-Line Treatment of Advanced Renal Cell Carcinoma,” February 28, 2011. <https://www.nice.org.uk/guidance/TA219/documents/renal-cell-carcinoma-second-line-metastatic-everolimus-appeal-decision2>.
- . “Decision of the Appeal Panel: Appraisal of the Use of Immunosuppressive Therapy for Renal Transplantation,” April 26, 2004. <https://www.nice.org.uk/guidance/ta85/documents/appraisal-of-the-use-of-immunosuppressive-therapy-for-renal-transplantation-decision-of-the-panel-2>.
- . “Decision of the Appeal Panel: Appraisal of the Use of Photodynamic Therapy for Age-Related Macular Degeneration,” March 17, 2003. <https://www.nice.org.uk/guidance/TA68/documents/decision-of-the-appeal-panel-appraisal-of-the-use-of-photodynamic-therapy-for-agerelated-macular-degeneration-2>.
- . “Decision of the Appeal Panel: Appraisal on the Use of Anakinra for the Treatment of Rheumatoid Arthritis,” September 18, 2003. <https://www.nice.org.uk/guidance/ta72/documents/appraisal-of-the-use-of-anakinra-for-the-treatment-of-rheumatoid-arthritis-decision-of-the-appeal-panel2>.
- . “Decision of the Appeal Panel: Donepezil, Rivastigmine, Galantamine and Memantine for the Treatment of Alzheimer’s Disease, Appeal by the Alzheimer’s Society, Age Concern, Counsel and Care, the Dementia Care Trust and the Royal College of Nursing,” July 14, 2006. <http://www.nice.org.uk/guidance/TA111/documents/appeal-by-the-alzheimers->

society-age-concern-counsel-and-care-the-dementia-care-trust-and-the-royal-college-of-nursing2.

- . “Decision of the Appeal Panel: Guidance on Erlotinib for the Treatment of Non-Small Cell Lung Cancer,” June 6, 2007. <https://www.nice.org.uk/guidance/ta162/documents/appeal-decision2>.
- . “Drugs for Alzheimer’s Disease: NICE Response to Consultee and Commentator Comments on the ACD-2 (January 2006),” May 2006. <https://www.nice.org.uk/guidance/ta111/documents/alzheimers-disease-responses-to-consultee-and-commentator-comments2>.
- . “Erlotinib for the Treatment of Non-Small-Cell Lung Cancer.” *Guidance*, November 2008. <https://www.nice.org.uk/guidance/ta162>.
- . “Etanercept and Efalizumab for the Treatment of Adults with Psoriasis: Post-Appeal Final Appraisal Determination (TA103),” June 9, 2006. <http://www.nice.org.uk/guidance/TA103/documents/postappeal-final-appraisal-determination-psoriasis-efalizumab-and-etanercept>.
- . “Etanercept and Infliximab for the Treatment of Psoriatic Arthritis.” *Guidance*, July 2006. <https://www.nice.org.uk/guidance/ta104>.
- . “Final Appraisal Determination: Everolimus for the Second-Line Treatment of Advanced Renal Cell Carcinoma,” June 2010. <https://www.nice.org.uk/guidance/ta219/documents/renal-cell-carcinoma-second-line-metastatic-everolimus-final-appraisal-determination3>.
- . “Guide to the Methods of Technology Appraisal 2013.” NICE, April 2013. <https://www.nice.org.uk/article/PMG9/>.
- . *Guide to the Processes of Technology Appraisal*. London: NICE, 2014. <http://www.nice.org.uk/article/pmg19/>.

- . “Guide to the Processes of Technology Appraisal.” NICE, September 2014. <https://www.nice.org.uk/article/pmg19/>.
- . “Guide to the Technology Appraisal and Highly Specialised Technologies Appeal Process,” February 2014. <https://www.nice.org.uk/article/PMG18/chapter/Foreword>.
- . “Methods for the Development of NICE Public Health Guidance (third Edition),” September 2012. <https://www.nice.org.uk/article/pmg4/chapter/reviewing-the-scientific-evidence>.
- . “NICE Charter,” April 2013. https://www.nice.org.uk/Media/Default/About/Who-we-are/NICE_Charter.pdf.
- . “NICE Statistics.” Accessed August 18, 2015. <https://www.nice.org.uk/news/nice-statistics>.
- . “NICE Technology Appraisal Guidance.” Accessed August 13, 2015. <https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/nice-technology-appraisal-guidance>.
- . “Public Involvement Programme.” *Public Involvement*. Accessed August 13, 2015. <https://www.nice.org.uk/about/nice-communities/public-involvement/public-involvement-programme>.
- . “Putting Patients and the Public at the Heart of NICE’s Work.” *Public Involvement*. Accessed August 13, 2015. <https://www.nice.org.uk/about/nice-communities/public-involvement>.
- . *Social Value Judgements: Principles for the Development of NICE Guidance*. 2nd ed. London: NICE, 2008. <https://www.nice.org.uk/proxy/?sourceurl=http://www.nice.org.uk/aboutnice/howwework/socialvaluejudgements/socialvaluejudgements.jsp>.

- . “Who We Are.” *About Us*. Accessed August 13, 2015. <https://www.nice.org.uk/about/who-we-are>.
- NHS Choices. “PIP Breast Implants - NHS Choices,” September 7, 2014. <http://www.nhs.uk/Conditions/Breast-implants/Pages/PIP-introduction.aspx>.
- NICE Public Involvement Programme. “Developing Technnnology Appraisal Guidance: A Factsheet for Patient and Carer Groups,” April 2014. <https://www.nice.org.uk/Media/Default/About/NICE-Communities/Public-involvement/Developing-NICE-guidance/Developing-technology-appraisals-factsheet-patient-carer-groups.pdf>.
- . “Hints and Tips for Nominating Patient Experts for a Single Technology Appraisal (STA),” July 2014. <http://www.nice.org.uk/Media/Default/About/NICE-Communities/Public-involvement/Developing-NICE-guidance/sta-tips-nominating-patient-experts.pdf>.
- . “Overview of Technology Appraisal: A Factsheet for Patient and Carer Organisations,” April 2014. <https://www.nice.org.uk/Media/Default/About/NICE-Communities/Public-involvement/Developing-NICE-guidance/Overview-of-technology-appraisals-patient-carer-groups.pdf>.
- . “Scoping Technology Appraisals: A Factsheet for Patient and Carer Organisations,” April 2014. <https://www.nice.org.uk/Media/Default/About/NICE-Communities/Public-involvement/Developing-NICE-guidance/Scoping-technology-appraisals-factsheet-patient-carer-organisations.pdf>.
- Norris, Sonya, and Marlisa Tiedemann. “Legal Status at the Federal Level of Assisted Human Reproduction in Canada.” Library of Parliament Background Paper. Ottawa: Library of Parliament, April 9, 2015.

- Nuffield Council on Bioethics. "Donor Conception: Ethical Aspects of Information Sharing." London: Nuffield Council on Bioethics, April 2013. <http://nuffieldbioethics.org/project/donor-conception/>.
- . "Emerging Biotechnologies: Technology, Choice and the Public Good." London: Nuffield Council on Bioethics, December 2012. <http://nuffieldbioethics.org/project/emerging-biotechnologies/>.
- . "Novel Neurotechnology: Intervening in the Brain." London: Nuffield Council on Bioethics, June 2013. <http://nuffieldbioethics.org/project/neurotechnology/>.
- OCEBM Levels of Evidence Working Group. "OCEBM Levels of Evidence." Oxford Centre for Evidence-Based Medicine. Accessed July 17, 2015. <http://www.cebm.net/index.aspx?o=5653>.
- O'Donnell, Guillermo A. "Horizontal Accountability in New Democracies." *Journal of Democracy* 9, no. 3 (1998): 112–26. doi:10.1353/jod.1998.0051.
- O'Mullan, Cathy, Maryanne Doherty, Rosemary Coates, and P. J. Matt Tilley. "Women's Experiences of Coping with the Sexual Side Effects of Antidepressant Medication." *Psychology & Health* 29, no. 12 (2014): 1388–1406. doi:10.1080/08870446.2014.940951.
- O'Neill, Onora. *A Question of Trust: The BBC Reith Lectures 2002*. Cambridge: Cambridge University Press, 2002.
- Parfit, Derek. *Reasons and Persons*. Oxford University Press, 1984.
- Park, Alison, Roger Jowell, and Suzi Mcpherson Contents. *The Future of the National Health Service: Results from a Deliberative Poll*, 1998.
- Pauline Quennell. "Getting a Word in Edgeways? Patient Group Participation in the Appraisal Process of the National Institute for Clinical Excellence." *Clinical*

- Governance: An International Journal* 8, no. 1 (March 1, 2003): 39–45.
doi:10.1108/14777270310459968.
- . “Getting Their Say, or Getting Their Way? - Has Participation Strengthened the Patient ‘voice’ in the National Institute for Clinical Excellence?” *Journal of Management in Medicine* 15, no. 3 (June 1, 2001): 202–19.
doi:10.1108/EUM0000000005509.
- Perry, Michael J. *Love and Power: The Role of Religion and Morality in American Politics*. Oxford: Oxford University Press, 1993.
- Persad, Govind, Alan Wertheimer, and Ezekiel J. Emanuel. “Principles for Allocation of Scarce Medical Interventions.” *Lancet (London, England)* 373, no. 9661 (January 31, 2009): 423–31. doi:10.1016/S0140-6736(09)60137-9.
- Petticrew, Mark, and Helen Roberts. “Evidence, Hierarchies, and Typologies: Horses for Courses.” *Journal of Epidemiology and Community Health* 57, no. 7 (July 2003): 527–29. doi:10.1136/jech.57.7.527.
- Pettit, Philip. “Depoliticizing Democracy.” *Ratio Juris* 17, no. 1 (March 1, 2004): 52–65.
doi:10.1111/j.0952-1917.2004.00254.x.
- Presthus, Robert. “Mrs. Thatcher Stalks the Quango: A Note on Patronage and Justice in Britain.” *Public Administration Review* 41, no. 3 (May 1, 1981): 312–17.
doi:10.2307/975791.
- Prior, Lindsay. “Belief, Knowledge and Expertise: The Emergence of the Lay Expert in Medical Sociology.” *Sociology of Health & Illness* 25, no. 3 (April 1, 2003): 41–57. doi:10.1111/1467-9566.00339.
- Public Administration Select Committee. *Smaller Government: Shrinking the Quango State*. 7 January 2011. HC 537-I 2010-2011.

- Quinn, Philip L. "Political Liberalisms and Their Exclusions of the Religious." *Proceedings and Addresses of the American Philosophical Association* 69, no. 2 (November 1, 1995): 35–56. doi:10.2307/3130495.
- Quong, Jonathan. *Liberalism Without Perfection*. Oxford; New York: OUP Oxford, 2010.
- Ramesh, Randeep. "NHS Postcode Lottery Survey Reveals Wide UK Disparities." *The Guardian*, December 9, 2011, sec. Society. <http://www.theguardian.com/society/2011/dec/09/nhs-lottery-survey-uk-disparities>.
- Rasmussen, Lisa M. "An Ethics Expertise for Clinical Ethics Consultation." *The Journal of Law, Medicine & Ethics* 39, no. 4 (2011): 649–61. doi:10.1111/j.1748-720X.2011.00632.x.
- . "Engineering, Gerrymandering and Expertise in Public Bioethics." *HEC Forum* 18, no. 2 (June 1, 2006): 125–30. doi:10.1007/s10730-006-9002-5.
- Rawlins, Michael. "Background to NICE's Citizens Council." In *Patients, the Public, and Priorities in Healthcare*, edited by Michael Rawlins and Peter Littlejohns. Oxford: Radcliffe Publishing, 2009.
- . "Background to NICE's Citizens Council." In *Patients, the Public and Priorities in Healthcare*, edited by Peter Littlejohns and Michael Rawlins. Oxford: Radcliffe Publishing, 2009.
- Rawlins, Michael D. "Pharmacopolitics and Deliberative Democracy." *Clinical Medicine (London, England)* 5, no. 5 (October 2005): 471–75.
- Rawls, John. *Political Liberalism: Expanded Edition*. 2nd ed. Columbia University Press, 2005.
- . "Political Liberalism: Reply to Habermas." *The Journal of Philosophy* 92, no. 3 (March 1, 1995): 132–80. doi:10.2307/2940843.

- Reinhardt, U. E. "Rationing Health Care: What It Is, What It Is Not, and Why We Cannot Avoid It." *The Baxter Health Policy Review* 2 (1996): 63–99.
- "Report of the Committee of Inquiry into Human Fertilisation and Embryology." Cmnd 9314, July 1984. http://www.hfea.gov.uk/docs/Warnock_Report_of_the_Committee_of_Inquiry_into_Human_Fertilisation_and_Embryology_1984.pdf.
- Robertson, John A. "Extending Preimplantation Genetic Diagnosis: The Ethical Debate Ethical Issues in New Uses of Preimplantation Genetic Diagnosis." *Human Reproduction* 18, no. 3 (March 1, 2003): 465–71. doi:10.1093/humrep/deg100.
- Robinson, L., D. Hutchings, L. Corner, T. Finch, J. Hughes, K. Brittain, and J. Bond. "Balancing Rights and Risks: Conflicting Perspectives in the Management of Wandering in Dementia." *Health, Risk & Society* 9, no. 4 (December 1, 2007): 389–406. doi:10.1080/13698570701612774.
- Rycroft-Malone, Jo, Kate Seers, Angie Titchen, Gill Harvey, Alison Kitson, and Brendan McCormack. "What Counts as Evidence in Evidence-Based Practice?" *Journal of Advanced Nursing* 47, no. 1 (July 2004): 81–90. doi:10.1111/j.1365-2648.2004.03068.x.
- Salter, Brian. "Who Rules? The New Politics of Medical Regulation." *Social Science & Medicine* 52, no. 6 (March 2001): 871–83. doi:10.1016/S0277-9536(00)00190-8.
- Sander, Jeffry D., and J. Keith Joung. "CRISPR-Cas Systems for Editing, Regulating and Targeting Genomes." *Nature Biotechnology* 32, no. 4 (April 2014): 347–55. doi:10.1038/nbt.2842.
- Savulescu, Julian. "Deaf Lesbians, 'designer Disability,' and the Future of Medicine." *BMJ* 325, no. 7367 (October 5, 2002): 771–73. doi:10.1136/bmj.325.7367.771.
- Saward, Michael. "Rawls and Deliberative Democracy." In *Democracy as Public Deliberation: New Perspectives*, edited by Maurizio Passerin D'Entreves, 112–

30. Manchester: Manchester University Press, 2002.
http://www.manchesteruniversitypress.co.uk/information_areas/subjects/politics/Perspectives_Democratization.htm.
- Schicktanz, Silke, and Michael Dusche. "The Ethics of Ethical Expertise in Science, Medicine and Healthcare Policies." *Indian Journal of Medical Ethics* 8, no. 3 (2011). <http://www.ijme.in/index.php/ijme/article/view/223/938>.
- Schlender, M. "The Use of Cost-Effectiveness by the National Institute for Health and Clinical Excellence (NICE): No(t Yet An) Exemplar of a Deliberative Process." *Journal of Medical Ethics* 34, no. 7 (July 1, 2008): 534–39. doi:10.2307/27720131.
- Schwartzman, Micah. "The Completeness of Public Reason." *Politics, Philosophy & Economics* 3, no. 2 (May 1, 2004): 191–220. doi:10.1177/1470594X04042963.
- Science and Technology Select Committee. *Developments in Human Genetics and Embryology*. 10 July 2002. HC 791 2001-2002, Ev 1.
- Science and Technology Select Committee. *Evidence Check 2: Homeopathy*. 8 February 2010. HC 45 2009-2010.
- Science and Technology Select Committee. *Human Reproductive Technologies and the Law*. 14 March 2005. HC 7-I 2004-2005.
- Shah, Koonal K. "Severity of Illness and Priority Setting in Healthcare: A Review of the Literature." *Health Policy (Amsterdam, Netherlands)* 93, no. 2–3 (December 2009): 77–84. doi:10.1016/j.healthpol.2009.08.005.
- Sher, George. "Subsidized Abortion: Moral Rights and Moral Compromise." *Philosophy & Public Affairs* 10, no. 4 (October 1, 1981): 361–72. doi:10.2307/2265049.
- Singer, Peter. "Moral Experts." *Analysis* 32, no. 4 (1972): 115–17.

- Skene, Loane. "Bioscience, Community Expectations and the Law." *The University of New South Wales Law Journal* 26, no. 3 (2003): 799–806.
- Skene, Loane, Ian Kerridge, Barry Marshall, Pamela McCombe, and Peter Schofield. "The Lockhart Committee: Developing Policy through Commitment to Moral Values, Community and Democratic Processes." *Journal of Law and Medicine* 16, no. 1 (August 2008): 132–38.
- Steinkamp, Norbert L., Bert Gordijn, and Henk A. M. J. ten Have. "Debating Ethical Expertise." *Kennedy Institute of Ethics Journal* 18, no. 2 (2008): 173–92.
- Stephenson, Professor Terence. "Expert Clinical Advice - MHRA Medical Devices: Report of the Independent Review on MHRA Access to Clinical Advice and Engagement with the Clinical Community in Relation to Medical Devices," July 18, 2014. <http://webarchive.nationalarchives.gov.uk/20141205150130/http://www.mhra.gov.uk/home/groups/commspo/documents/news/con402542.pdf>.
- Stewart, Moira. "Towards a Global Definition of Patient Centred Care." *BMJ: British Medical Journal* 322, no. 7284 (February 24, 2001): 444–45.
- Syrett, Keith. "Deconstructing Deliberation in the Appraisal of Medical Technologies: NICEly Does It?" *The Modern Law Review* 69, no. 6 (November 1, 2006): 869–94. doi:10.1111/j.1468-2230.2006.00615.x.
- Taylor, Rod S., and Cynthia P. Iglesias. "Assessing the Clinical and Cost-Effectiveness of Medical Devices and Drugs: Are They That Different?" *Value in Health* 12, no. 4 (June 1, 2009): 404–6. doi:10.1111/j.1524-4733.2008.00476_2.x.
- Tenbensel, Tim. "Virtual Special Issue Introduction: Public Participation in Health Policy in High Income Countries – A Review of Why, Who, What, Which, and Where?" *Social Science & Medicine* 71, no. 9 (November 2010): 1537–40. doi:10.1016/j.socscimed.2010.08.005.

- Thatcher, Mark, and Alec Stone Sweet. "Theory and Practice of Delegation to Non-Majoritarian Institutions." *West European Politics* 25, no. 1 (January 1, 2002): 1–22. doi:10.1080/713601583.
- The General Product Safety Regulations 2005*, n.d. <http://www.legislation.gov.uk/uksi/2005/1803/contents/made>.
- The Human Fertilisation and Embryology Act of 1990*, c. 37. <http://www.legislation.gov.uk/ukpga/1990/37/contents>.
- The Human Fertilisation and Embryology Act of 2008*, c. 22. <http://www.legislation.gov.uk/ukpga/2008/22/contents>.
- Thompson, Jill, Paul Bissell, Cindy Cooper, Chris J. Armitage, and Rosemary Barber. "Credibility and the 'professionalized' Lay Expert: Reflections on the Dilemmas and Opportunities of Public Involvement in Health Research." *Health*: 16, no. 6 (November 1, 2012): 602–18. doi:10.1177/1363459312441008.
- Thorne, Sally E, Kerstin Ternulf Nyhlin, and Barbara L Paterson. "Attitudes toward Patient Expertise in Chronic Illness." *International Journal of Nursing Studies* 37, no. 4 (August 1, 2000): 303–11. doi:10.1016/S0020-7489(00)00007-9.
- Thraves, Laurie. "The Future of Arm's-Length Bodies." London: Local Government Information Unit, March 19, 2012. <http://www.lgiu.org.uk/wp-content/uploads/2012/04/Future-of-arms-length-bodies.pdf>.
- "Tories Pledge to Cut Back Quangos." *BBC News*, July 6, 2009. http://news.bbc.co.uk/go/pr/fr/-/1/hi/uk_politics/8135460.stm.
- Tran, Joseph, Rosanna Tran, and John R. White. "Smartphone-Based Glucose Monitors and Applications in the Management of Diabetes: An Overview of 10 Salient 'Apps' and a Novel Smartphone-Connected Blood Glucose Monitor." *Clinical Diabetes* 30, no. 4 (October 1, 2012): 173–78. doi:10.2337/diaclin.30.4.173.

- UK BioIndustry Association, and Medicines and Healthcare Products Regulatory Agency. "Pathway of Innovation from Research to Patients Conference." London, June 11, 2015. http://www.biamhraconference.org/Conference_report_04_web.pdf.
- Upshur, R.E.G., E.G. Vandenberg, and V. Goel. "Meaning and Measurement: An Inclusive Model of Evidence in Health Care." *Journal of Evaluation in Clinical Practice* 7, no. 2 (2001): 91–96. doi:10.1046/j.1365-2753.2001.00279.x.
- Waller, Louis. "Reproductive Technology: In Australia, the Debate Moves to Embryo Experimentation." *The Hastings Center Report* 17, no. 3 (June 1987): S21–22.
- Walters, LeRoy. "Ethics and New Reproductive Technologies: An International Review of Committee Statements." *The Hastings Center Report* 17, no. 3 (June 1, 1987): 3–9. doi:10.2307/3562251.
- Weijer, Charles. "The Ethical Analysis of Risk." *The Journal of Law, Medicine & Ethics* 28, no. 4 (December 1, 2000): 344–61. doi:10.1111/j.1748-720X.2000.tb00686.x.
- Welsh, S., A. Hassiotis, G. O'Mahoney, and M. Deahl. "Big Brother Is Watching You--the Ethical Implications of Electronic Surveillance Measures in the Elderly with Dementia and in Adults with Learning Difficulties." *Aging & Mental Health* 7, no. 5 (September 2003): 372–75. doi:10.1080/1360786031000150658.
- "What's the Issue with Mesh Implants?" *BBC News*, June 17, 2014. <http://www.bbc.co.uk/news/uk-scotland-27887766>.
- Wilkinson, Dominic, G. Owen Schaefer, Kelton Tremellen, and Julian Savulescu. "Double Trouble: Should Double Embryo Transfer Be Banned?" *Theoretical Medicine and Bioethics* 36, no. 2 (April 2015): 121–39. doi:10.1007/s11017-015-9324-x.

- Wilkinson, Stephen, and Eve Garrard. "Mind the Gap: The Use of Empirical Evidence in Bioethics." In *Bioethics and Social Reality*, edited by Matti Häyry, Tuija Takala, and Peter Herissone-Kelly, 73–87. Value Inquiry Book Series 165. New York: Rodopi bv, 2005.
- Wilmshurst, Peter. "The Regulation of Medical Devices." *BMJ* 342 (May 14, 2011): d2822. doi:10.1136/bmj.d2822.
- Wilson, Duncan. "Creating the 'ethics Industry': Mary Warnock, in Vitro Fertilization and the History of Bioethics in Britain." *Biosocieties* 6, no. 2 (June 2011): 121–41. doi:10.1057/biosoc.2010.26.
- . *The Making of British Bioethics*. Manchester (UK): Manchester University Press, 2014. <http://www.ncbi.nlm.nih.gov/books/NBK247179/>.
- Wong, David B. "Coping with Moral Conflict and Ambiguity." *Ethics* 102, no. 4 (July 1, 1992): 763–84. doi:10.2307/2381553.