DECLARATIONS OF UNCONSTITUTIONALITY IN THE COMMON LAW TRADITION: A COMPARATIVE AND THEORETICAL ANALYSIS

ROBERT NOONAN
Trinity College
University of Dublin

Supervisor: Dr. Oran Doyle

School of Law
Faculty of Arts, Humanities and Social Sciences

A dissertation submitted in partial fulfilment of the requirements of the degree of
Doctor of Philosophy

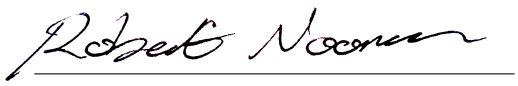
2019

Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work.

I agree to deposit this thesis in the University's open access institutional repository or allow the library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

Trinity College Dublin, 28 September 2018



Robert Noonan

Summary and Methods

The principal questions of this thesis are: (1) how is unconstitutionality practised and theorised in the common law tradition, and (2) whether improvements can be made to this theory and practice. Constitutional orders in which judges have a full power to review the compatibility of legislation with the constitution must invariably confront the difficulty of unconstitutional law. To investigate this issue, the thesis combines a comparative legal analysis of the laws of Ireland, Canada, the United States, India and South Africa with a jurisprudential analysis of philosophical issues that inform unconstitutionality as a general legal phenomenon.

Research Methods

The comparator jurisdictions under scrutiny in this thesis were selected on the basis of several features: they each mainly practice strong-form judicial review against the institutional background of a decentralised court system with a preference for examining the constitutionality of law only as it applies to actual, concrete cases. They also largely share a common law heritage. The aim of maintaining these constants is to sharpen contrast where it appears in the practices of these legal systems. The primary objects of study are thus the constitutional texts and case law of the five jurisdictions outlined above.

The thesis also employs theoretical reflection to synthesise the data that is generated from the comparative work. Thus, theoretical literature on the nature of law and the nature of legal systems is also employed to assist in developing models to reconstruct the practice of unconstitutionality and to track and explain where deviations between the jurisdictions occur. Where these deviations are a result of theoretical choice, the reasons that might underlie this choice are also subjected to critical analysis, with a view to articulating an all-things-considered preferable theory of unconstitutionality.

Major Findings of the Thesis

This thesis is the first study to provide an account of the practice of unconstitutionality of such a broad base of jurisdictions with a common law heritage. It is also a significant attempt to systematically categorise unconstitutionality, at least as it occurs in these jurisdictions. To guide the formation of theoretical models of unconstitutionality, I argue that there are three questions a theory of unconstitutionality must answer: the *derivation question*, the *effects question*, and the *temporal question*. The derivation question accounts for the source of unconstitutionality and answers, from a more technocratic point of view, how unconstitutionality is brought about in the first place. The effects question addresses how unconstitutionality modifies the properties of legal norms. Finally, the temporal question attends to when unconstitutionality begins, and for how long it is effective.

Using these questions, I both present a scheme for reconstructing the practice of the studied jurisdictions, and I propose a new model of unconstitutionality. The comparative study identifies several trends and recurrent issues among the compar-

ator jurisdictions. In particular, there is a tendency, with some slight deviation, to find unconstitutional law invalid, and void *ab initio*. The thesis challenges this understanding of unconstitutionality, demonstrating how a more fine-grained theory of the nature of law and legal systems can provide an alternative understanding of unconstitutionality that both coheres with practice and mitigates the severity of the effects achieved by unconstitutionality. The thesis argues that validity is an important aspect of unconstitutionality, but that other properties of legal norms that are often ignored, such as validity, are perhaps even more important. Abandoning the all-or-nothing view of validity as the one controlling concept in unconstitutionality allows for greater flexibility in cases where legal norms are found unconstitutional.

The reconsideration of the role of validity produces several interesting observations and suggestions. One of these is that the suspended declaration of unconstitutionality should be thought of as regulating the applicability of unconstitutional norms, and not the timing of the invalidity of those norms. This resolves some otherwise difficult contradictions in the possibility of that practice at all in many jurisdictions (including those that have already endorsed it, such as Canada and South Africa). The abandonment of validity as the core, controlling concept in unconstitutionality also prompts greater transparency in the consideration of important value-judgments in constitutional adjudication. Given the close connection between constitutional law, politics, and morality, this transparency is arguably preferable over a validity analysis that effaces these values by subsuming them into a technical validity inquiry with a strictly binary result.

The thesis thus deepens knowledge of the law as it stands and also makes normative proposals for improving theory and practice in this area. The thesis' preferred model of unconstitutionality is developed on a 'choice of law' analogy that coheres with a detailed theory of legal systems endorsed by this thesis. This model is both theoretically and practically preferable to competitor models and, if adopted by courts, could produce new legal doctrines that deal more sensibly with the moral dimensions of retrospectively imposing changes to legal rules, providing much-needed clarity in this area of public law.

Acknowledgements

The research for this thesis was carried out under the auspices of Trinity College Dublin. I am very grateful for the support given to me by the university, in the form of both a non-foundation scholarship and an Ussher Fellowship. Without this generosity, it is highly unlikely I would have been able to undertake and complete this work.

I am grateful to my supervisor, Dr. Oran Doyle, for his guidance and encouragement. His exceptional talent for precise and sophisticated analysis was of immense benefit to me in receiving feedback on my work and inspiring me to further clarify and defend my arguments. From my induction into constitutional law as a first year undergraduate through to this thesis, I could have had no better guide. As Dante Alighieri's famous literary work reminds us; it is very important to have a good guide when embarking on a daunting venture such as a PhD thesis.

I wish also to thank my fellow postgraduate students in the School of Law for many hours of interesting conversations, in particular David Campbell and Tom Kelly. Postgraduate legal research is often acknowledged to be a lonely affair, and I was very lucky to have such stalwart companions.

My colleagues at the Law Reform Commission of Ireland have also been exceptionally tolerant in attentively listening to, and insightfully critiquing, some of my ideas in this thesis. In particular, I am grateful to Commissioner Ray Byrne for being so generous with his time and for repeatedly pressing me on ideas and compelling me to improve them.

I would also like to express my gratitude to Ezequiel Monti, with whom I shared many conversations on philosophical topics during my masters degree at Oxford. Ezequiel's incredible insight and perception no doubt made me a better reasoner. He also very generiously gifted me a copy of Eugenio Bulygin's *Essays in Legal Philosophy* a work to which, as will become apparent, I owe a significant intellectual debt and from which I continue to learn.

Finally, and most importantly of all, I owe particular thanks to Lisa Harpur for all her love and support throughout this project. The efforts of everyone above would have been set at naught were it not for her inexhaustible compassion.

I have endeavoured to state the law as of 28 September 2018. Any errors and omissions are my own.

Table of Contents

DECLARATION i

SUMMARY AND METHODS ii

ACKNOWLEDGEMENTS v

TABLE OF CASES xv

Ireland xv

England and Wales xvii

Australia xvii

Canada xvii

India xix

The United States xx

South Africa xxiv

Other Jurisdictions xxvi

TABLE OF LEGISLATION xxvii

Ireland xxvii

England and Wales xxvii

South Africa xxvii

TABLE OF CONSTITUTIONAL PROVISIONS xxix

Ireland xxix

Canada xxix

United States xxix

India xxx

European Union xxx

South Africa xxx
Other Constitutions xxxi

LIST OF FIGURES xxxiii

LIST OF TABLES xxxv

PART I INTRODUCTION 1

1. INTRODUCTION 3

1.1. Subject of the Thesis 4
1.1.1. The Problem of Unconstitutionality 6

1.2. Methodology and Limitations 7
1.2.1. Functionalism and Universalism 8
1.2.2. Situating the Thesis in Comparative Scholarship 9
1.2.3. Jurisdiction Selection 11
1.2.3.1. Strong and Weak Judicial Review 12
1.2.3.2. Centralised and Decentralised Courts 13
1.2.3.3. Concrete and Abstract Judicial Review 15
1.2.3.4. Jurisdictions Selected for Study & Limitations of the Thesis 16
1.2.3.5. Geographic Dispersal and Grounding of Judicial Review 18
1.2.4. The Role of Legal Theory 19

1.3. Structure of the Thesis 20
1.3.1. Part One: Introduction 20
1.3.2. Part Two: Declarations of Unconstitutionality in Practice . 20
1.3.3. Part Three: Theoretical Analysis 21
1.3.4. Part Four: Conclusion 23

2. CONSTITUTIONAL BASES FOR JUDICIAL REVIEW 25

2.1. Introduction 26

2.2. Textual Foundations 27
2.2.1. Concrete Review 27
2.2.1.1. Fundamental Rights Review 27
2.2.1.2. Federal Review 34
2.2.1.3. Observations 37

2.2.2. Abstract Review . . . 39
2.2.2.1. *Ireland* . . . 39
2.2.2.2. *Canada* . . . 41
2.2.2.3. *South Africa* . . . 41
2.2.2.4. *India* . . . 42
2.2.3. Constitutional Conditions for a Bill to Count as Law . . . 42
2.2.3.1. *United States* . . . 43
2.2.3.2. *Ireland* . . . 43
2.2.3.3. *Canada* . . . 44
2.2.3.4. *India* . . . 45
2.2.3.5. *South Africa* . . . 46
2.3. Non-Textual Foundations . . . 47
2.4. Conclusions . . . 50

PART II DECLARATIONS OF UNCONSTITUTIONALITY IN PRACTICE 53

3. THE UNITED STATES 55
3.1. Introduction . . . 56
3.2. Doctrines Avoiding or Curtailing Unconstitutionality . . . 57
3.2.1. Avoidance . . . 57
3.2.2. Reading In . . . 58
3.2.3. Severance . . . 59
3.2.4. *Res Judicata* . . . 61
3.3. Early Case Law on Unconstitutionality: Voidness *Ab Initio* . . . 62
3.4. Prospectivity in the Warren Court . . . 67
3.4.1. *Linkletter* and Prospective Invalidity in Criminal Cases . . . 67
3.4.2. *Chevron Oil* and Prospective Invalidity in Civil Cases . . . 75
3.5. The Modern Approach to Unconstitutionality and Retroactivity . . . 80
3.5.1. Retreating from the *Marbury/Norton* View . . . 80
3.5.2. The New Rules on Retrospectivity . . . 83
3.5.2.1. *Criminal Cases on Direct Review* . . . 83
3.5.2.2. *Habeas Corpus Applications* . . . 84
3.5.2.3. *Federal Civil Disputes* . . . 85
3.6. As-Applied and Facial Challenges . . . 87
3.7. Revival of Unconstitutional Statutes . . . 94

3.8. **Conclusion** 97

4. CANADA **101**

4.1. **Introduction** 101

4.2. **Doctrines Avoiding or Curtailing Unconstitutionality** 102

4.2.1. The Presumption of Constitutionality 103

4.2.2. Reading In 104

4.2.3. Severance 105

4.3. **Early Case Law: Unconstitutionality as Voidness/Invalidity** . . . 105

4.3.1. Restitution for Unconstitutional Taxation 106

4.3.2. Review under the Charter: *Big M Drug Mart* 108

4.3.3. The *Manitoba Language Rights* Reference 110

4.4. **The Rise of the Suspended Declaration** 113

4.4.1. *Schachter v Canada* 113

4.4.2. The Proliferation of Suspended Declarations: *Bedford* and *Carter* 114

4.5. **The Revival of Unconstitutional Law** 116

4.6. **Conclusion** 117

5. IRELAND **119**

5.1. **Introduction** 119

5.2. **Doctrines Avoiding or Curtailing Unconstitutionality** 120

5.2.1. The Presumption of Constitutionality and Double Construction 120

5.2.2. Severance 121

5.3. **Unconstitutionality in the Irish Courts** 122

5.3.1. Early Case Law: *Murphy v Attorney General* 122

5.3.2. Drifting from the *Murphy* Standard 125

5.3.3. Completing the Drift: *A v Governor of Arbour Hill Prison* . 128

5.4. **The Modern Position on Declarations of Unconstitutionality** . . 131

5.5. **Suspended Declarations of Unconstitutionality** 134

5.6. **Pre-Constitution Law** 136

5.7. **Revival of Unconstitutional Law** 139

5.8. **Conclusion** 141

6. INDIA 143
6.1. Introduction . . . 143
6.2. Doctrines Avoiding or Curtailing Unconstitutionality . . . 144
6.2.1. Presumption of Constitutionality and Double Construction 144
6.2.2. Severance . . . 144
6.3. Degrees of Voidness . . . 145
6.3.1. Voidness under Article 13(1) . . . 147
6.3.2. Voidness under Article 13(2) . . . 153
6.3.3. Some Clarity: *Jagannath* and *Ambica Mills* . . . 154
6.4. The Doctrine of Eclipse . . . 157
6.4.1. Extension of the Doctrine to Post-Constitution Law? . . . 160
6.4.2. Restrictions on the Doctrine of Eclipse . . . 161
6.5. Conclusion . . . 163
7. SOUTH AFRICA 165
7.1. Introduction . . . 165
7.2. Doctrines Avoiding or Curtailing Unconstitutionality . . . 166
7.2.1. Reading In . . . 166
7.2.2. Severance . . . 168
7.2.2.1. Actual Severance . . . 169
7.2.2.2. Notional Severance . . . 170
7.3. The Doctrine of 'Objective Invalidity' . . . 171
7.3.1. The Basis of the Doctrine: *Ferreira v Levin* . . . 172
7.3.2. Recent Applications of the Doctrine . . . 174
7.3.3. *Erga Omnes* Effect . . . 177
7.4. Suspended Declarations of Unconstitutionality . . . 178
7.4.1. Constitutional Basis and Frequency of Suspension . . . 179
7.4.2. Interim Orders . . . 181
7.5. Conclusion . . . 182
PART III THEORETICAL ANALYSIS 183
8. A THEORETICAL FRAMEWORK FOR UNCONSTITUTIONALITY 185

8.1. Introduction . . . 186
8.2. Three Questions for a Theory of Unconstitutionality . . . 188
8.3. The Derivation of Unconstitutionality . . . 190
8.3.1. Judicial Intervention . . . 192
8.3.1.1. Kelsen's Argument . . . 192
8.3.1.2. Legal Realism . . . 196
8.3.2. Legislative Nullity . . . 198
8.3.2.1. Blackstone's Declaratory Theory . . . 202
8.3.2.2. 'Right Answer' Theories . . . 204
8.3.2.3. The Role of Interpretation . . . 210
8.3.3. Differentiating Legislative Nullity and Judicial Intervention 211
8.4. The Effect of Unconstitutionality . . . 215
8.4.1. Validity . . . 216
8.4.2. Applicability . . . 221
8.4.3. Invalidity does not Imply Inapplicability . . . 225
8.4.3.1. D-Validity ≠ Applicability . . . 227
8.4.3.2. N-Validity ≠ Applicability . . . 231
8.4.4. Summary and Conclusion . . . 232
8.5. The Timing of Unconstitutionality . . . 235
8.5.1. The Origin Question . . . 235
8.5.1.1. Unconstitutionality from the Future . . . 235
8.5.1.2. The Simulation of Retrospective Effect . . . 238
8.5.2. The Duration Question . . . 239
8.6. Combining Answers to the Three Questions . . . 240
8.6.1. Voidness *Ab Initio* (Ln | Va | Pa | Rt & Pr) . . . 243
8.6.2. Applicability-Regulating Mechanisms (XX | Ap | XX | Rt & Pr) . . . 246
8.6.2.1. Pre-Constitution Law in Ireland and India (Ln | Ap | Pa | Rt & Pr) . . . 246
8.6.2.2. Ambica Mills in India (Ln | Ap | Pa | Rt & Pr) . . . 249
8.6.2.3. Suspended Declarations of Unconstitutionality (Ln | Ap | Pa | Rt & Pr) . . . 249
8.6.2.4. As-Applied Challenges (Ji | Ap | Pr | Rt & Pr) . . . 250
8.6.3. Prospective-Only Mechanisms (XX | XX | XX | Pr) . . . 251
8.6.3.1. The Legislative Nullity Variant: Ireland (Ln | Va | Pa | Pr) . . . 251

8.6.3.2. The Judicial Intervention Variant: United States (Ji | Ap | Pr | Pr) 251

8.7. Conclusion 252

9. THE PLACE OF UNCONSTITUTIONAL LAW IN THE LEGAL SYSTEM **255**

9.1. Introduction 255

9.2. Law and Legal System 257

9.3. Legal Systems as Sets 259

9.3.1. Relating Laws to Other Laws 260

9.3.2. Relating Systems to Other Systems 262

9.4. Legal System and Legal Order 263

9.4.1. Membership vs Inclusion 266

9.4.2. Ordered and Unordered Sets 271

9.4.3. The 'Internal and External' Time of Legal Norms 273

9.5. The Consistency and Completeness of Legal Systems 274

9.5.1. Raz's Error 274

9.5.2. Bulygin on Consistency and Completeness 276

9.6. Conclusion 278

10. FLAWS IN EXISTING THEORIES OF UNCONSTITUTIONALITY **283**

10.1. Introduction 284

10.2. The Character of Law Problem 285

10.2.1. Metaphysical Objectivity in Law 286

10.2.2. Dworkin and 'Modest' Objectivity 287

10.2.2.1. The Impossibility of the Herculean Standard . . . 288

10.2.2.2. Dworkin on the Possibility of Judicial Disagreement 291

10.2.3. Moore and 'Strong' Objectivity 293

10.2.3.1. K-P Semantics 293

10.2.3.2. Criticism of K-P Semantics 295

10.2.4. Conclusions on Legal Objectivity 298

10.3. The Retrocausality Problem 299

10.3.1. Motivating Judicial Law-Making 300

10.3.2. Judicial Law-Making and Legislative Law-Making 303

10.3.3. Law-Making Power Implies Law-Derogating Power 305
10.3.4. The Tension between Invalidity and Judicial Law-Making . 306
10.3.5. Two Solutions for Retrocausality 307
10.3.5.1. Resolution One: Eternal Law 307
10.3.5.2. Resolution Two: 'Quantum Law' 309
10.3.6. Conclusion 312
10.4. Institutional/Moral Problems 313
10.4.1. General Justice at the Cost of Individual Injustice 314
10.4.2. Response of Other State Institutions to Suspension 316
10.4.3. The Rule of Law 318
10.5. Conclusion 320

PART IV CONCLUSION **323**

11. IMPLICATIONS FOR CONSTITUTIONAL THEORY AND PRACTICE **325**

11.1. Introduction 325
11.2. The Preferred Theory of Unconstitutionality 327
11.2.1. Judicial Power as the Ground of Unconstitutionality 327
11.2.2. Inapplicability as the Effect of Unconstitutionality 328
11.2.3. Moral Considerations and the Temporality of Unconstitutionality 331
11.3. The Insufficiency of a Remedies-Centred Analysis 334
11.3.1. The Pragmatic Approach 335
11.3.2. The Conceptual Approach 337
11.3.3. Understanding the Suspended Declaration 340
11.3.4. Roach's Two-Track Remedies Proposal 341
11.4. Unconstitutionality and the Comparative Method 343
11.4.1. The Use of Foreign Law 343
11.4.2. The Limits of Textual Constraints 345
11.5. Conclusion 346

BIBLIOGRAPHY **349**

Articles 349
Books 358

Contributions to Collections 361
Other Works 365

Table of Cases

Ireland

A v Governor of Arbour Hill Prison [2006] IESC 45, [2006] 4 IR 88 . 38, 127–133, 141, 197, 244, 251, 285, 300, 308
AB v The Clinical Director of St Loman's Hospital [2018] IECA 123 .... 135, 198
An Blascaod Mór Teoranta v Commissioners of Public Works (No 3) [2000] 1 IR 6 (SC) 121
An Blascaod Mór Teoranta v Commissioners of Public Works (No 3) [1998] IEHC 38 ..... 137, 138, 213
Article 26 and the Housing (Private Rented Dwellings) Bill 1981, In Re [1983] IR 181 (SC) 40, 138

Bederev v Ireland [2015] IECA 38 134
Blake v Attorney General [1982] IR 117 (SC) 121, 133, 134, 214
Brennan v Attorney General [1983] ILRM 449 (HC) 138, 172
Browne v Attorney General [1991] 2 IR 58 (HC) 214

Campus Oil v Minister for Industry and Energy (No 2) [1983] IR 88 (SC) 133
CC v Ireland [2006] IESC 33, [2006] 4 IR 1 ..... 7, 128, 141
Clarke v Governor of Mountjoy Prison [2016] IEHC 278 132
Cox v Ireland [1992] 2 IR 53 (SC) 125, 126
Crilly v T & J Farrington Ltd [2001] 3 IR 251 (SC) 128
Cummins v McCartan [2005] 3 IR 559 121

Damache v Director of Public Prosecutions [2012] IESC 11, [2012] 2 IR 266 .. 73, 131, 251, 313
Deaton v Attorney General [1963] IR 170 121
Desmond v Glackin (No 2) [1993] 3 IR 67 (SC) 122

East Donegal Co-Operative Livestock Market v Attorney General [1970] IR 317 (SC) 120
ESB v Gormley [1985] IR 129 (SC) 137

Gaffney v Gaffney [1975] IR 133 (SC) 31
Geoghegan v Institute of Chartered Accountants in Ireland [1995] 3 IR 86 (HC) . 31, 137, 247
Glavin v Governor of Mountjoy Prison [1991] 2 IR 421 (SC) 130

Grealis v Director of Public Prosecutions [2001] 3 IR 144 (SC) 121, 131

Haughey, In Re [1971] IR 217 (SC) 120
Hegarty v O'Loughran [1990] 1 IR 148 121

JP v Director of Public Prosecutions [2009] 3 IR 215 (HC) 137, 247

Kennedy v Governor of Mountjoy Prison [2017] IEHC 402 132
Kinsella v Governor of Mountjoy Prison [2011] IEHC 235, [2012] 1 IR 467 .. 134
Kovacs v Governor of Mountjoy Women's Prison (Dochas Centre) [2015] IEHC 418 132

Laurentiu v Minister for Justice, Equality and Law Reform [1999] 4 IR 26 (SC) 303

Maher v Attorney General [1973] IR 140 (SC) 121, 138, 161
McDonald v Bord na gCon [1965] IR 217 (SC) 120
McDonnell v Ireland [1998] 1 IR 134 (SC) 125–129, 131, 133
McGee v Attorney General [1974] IR 284 (SC) 128, 214
McKinley v Minister for Defence [1992] 2 IR 333 (HC) 31
Muckley v Ireland [1985] IR 472 (SC) 125, 127, 128, 130, 131, 133
Murphy v Attorney General [1982] IR 241 (SC) . 38, 122–131, 133, 138, 141, 243, 244, 251

NHV v Minister for Justice [2017] IESC 82 135
NHV v Minister for Justice [2017] IESC 35 ..... 135, 141

O'Brien v Manufacturing Engineering Co Ltd [1973] IR 334 (HC) 137
O'Keeffe v An Bord Pleanála [1993] 1 IR 39 (SC) 133
Osinuga and Agha v Minister for Social Protection [2018] IECA 155 .... 135, 136

PC v Minister for Social Protection [2017] IESC 63 135
People (Attorney General) v Conmey [1975] IR 341 (SC) 137

Shelly v Mahon [1990] IR 36 (SC) 130

The People (DPP) v Cunningham [2012] IECCA 64, [2013] 2 IR 631 131
The People (DPP) v Doyle [2017] IESC 1 132
The People (DPP) v Hughes [2012] IECCA 69, [2013] 2 IR 619 131
The People (DPP) v Kavanagh [2012] IECCA 65 132
The People (DPP) v McKevitt [2014] IECCA 19 132
The People (DPP) v MS [2003] 1 IR 606 (SC) 31, 137, 247
The People (DPP) v O'Brien [2012] IECCA 68 132
The People (DPP) v Timmons [2013] IECCA 86 132
The People (DPP) v WM [1995] 1 IR 226 (HC) 140, 151
The State (Abenglen Properties Ltd) v Dublin Corporation [1984] IR 381 (SC) 133
The State (Byrne) v Frawley [1978] IR 326 (SC) 129
The State (Sheerin) v Kennedy [1966] IR 379 (SC) 31, 42, 137, 247
Thomas Hunter Ltd v James Fox and Co Ltd [1966] IR 520 (SC) 122

Thomondgate No 1 Area CPO, In Re 1986 WJSC-SC 1512 133

W v W [1992] 2 IR 476 (SC) 31
Willis v Governor of Wheatfield Prison [2015] IEHC 251 132

England and Wales

Cadder v Her Majesty's Advocate [2010] UKSC 43, [2010] 1 WLR 2601 128

Hoffmann-La Roche v Secretary of State for Trade [1975] AC 295 (HL) 178

Percy v Hall [1997] QB 924 (HC) 128

R v Governor of Brockhill Prison, ex parte Evans [2000] UKHL 48, [2001] 1 AC 19 128

Spectrum Plus Ltd (In Liquidation), In Re [2005] UHKL 41, [2005] 2 AC 680 202

Australia

Ha v New South Wales (1997) 189 CLR 465 (HC) 245

Canada

Air Canada v British Columbia 1989 CanLII 95 (SCC), [1989] 1 SCR 1161 .. 106, 107
Amax Potash v Government of Saskatchewan 1976 CanLII 15 (SCC), [1976] 2 SCR 576 106
Attorney General of Manitoba v Forest 1979 CanLII 242 (SCC), [1979] 2 SCR 1032 111

Bilodeau v Attorney General of Manitoba [1981] 5 WWR 393 (MCOA) 111

Canada (Attorney General) v Bedford 2013 SCC 72, [2013] 3 SCR 1101 . 115, 316
Canada (Attorney General) v Hislop 2007 SCC 10, [2007] 1 SCR 429 197, 202, 300
Carter v Canada (Attorney General) 2015 SCC 5, [2015] 1 SCR 331 .... 115, 303
Charkaoui v Canada (Citizenship and Immigration) [2007] 1 SCR 350 (SCC) 111

Dumore v Ontario 2001 SCC 94 (CanLII), [2001] 3 SCR 1016 114

Edwards v Canada (Attorney General) [1930] AC 124 (UKPC) 211
Eurig Estate, In Re [1998] 2 SCR 656 (SCC) 107, 111

Figueroa v Canada (Attorney General) [2003] 1 SCR 912 (SCC) 111, 114

Harper v Canada (Attorney General) 2000 SCC 57 (CanLII), [2000] 2 SCR 764 .. 103

Health Services and Support – Facilities Subsector Bargaining Association v British Colombia [2007] 2 SCR 391 (SCC) 111

Isaac v British Columbia (Superintendent of Motor Vehicles) 2014 BCSC 1608 (CanLII) 117

Kingstreet Investments Ltd v New Brunswick (Finance) 2007 SCC 1 (CanLII), [2007] 1 SCR 3 107

M v H [1999] 2 SCR 3 (SCC) 111
Mackin v New Brunswick [2002] 1 SCR 405 (SCC) 111
Manitoba (AG) v Metropolitan Stores Ltd 1987 CanLII 79 (SCC), [1987] 1 SCR 110 103
Manitoba Language Rights, In Re [1985] 1 SCR 721 (SCC) .. 7, 41, 110, 112, 113, 117, 285, 318
McCraken v McIntyre 1877 CanLII 16 (SCC), [1877] 1 SCR 479 109
McSorley v The Mayor of the City of St John 1882 CanLII 31 (SCC), [1882] 6 SCR 531 109
Miron v Trudel [1995] 2 SCR 418 (SCC) 105, 171

Nguyen v Quebéc (Education, Recreation and Sports) [2009] 3 SCR 208 (SCC) .. 112, 115
Nova Scotia (Workers' Compensation Board) v Martin; Nova Scotia (Workers' Compensation Board) v Laseur [2003] 2 SCR 504 (SCC) 109, 111

Ottawa Agriculture Insurance Co v Sheridan 1880 CanLII 21 (SCC), [1879] 5 SCR 157 109

R v Al-Qaysi 2016 BCSC 937 114
R v Big M Drug Mart Ltd [1985] 1 SCR 295 (SCC) 108, 110, 243, 245
R v Corporal SA Strong 2008 CM 3019 (CanLII) 109
R v Demers [2004] 2 SCR 489 (SCC) 111
R v DeSousa 1992 CanLII 80 (SCC), [1992] 2 SCR 944 246
R v Ferguson 2008 SCC 6 (CanLII), [2008] 1 SCR 96 246
R v Jordan 2016 SCC 27 (CanLII), [2016] 1 SCR 361 197
R v Mercure [1988] 1 SCR 234 (SCC) 112
R v Moazami 2014 BCSC 261 114
R v Nur 2015 SCC 15 (CanLII), [2015] 1 SCR 773 246
R v Paquette [1990] 2 SCR 1103 (SCC) 112
R v Rahey [1987] 1 SCR 588 (SC) 30
R v Swain 1991 CanLII 104 (SCC), [1991] 1 SCR 933 ..... 113, 135
R v Vader 2016 ABQB 309 (CanLII) 117
R v Wigman [1987] 1 SCR 246 (SCC) 230
R v Zundel 1992 CanLII 75 (SCC), [1992] 2 SCR 731 103
Ravndahl v Saskatchewan 2009 SCC 7 (CanLII), [2009] 1 SCR 181 116
Reference Re Same-Sex Marriage [2004] 3 SCR 698 (SCC) 303
Reference Re Secession of Quebec 1998 CanLII 793, [1998] 2 SCR 217 41

Reference Re Securities Act 2011 SCC 66, [2011] 3 SCR 837 41
Reference Re Senate Reform 2014 SCC 32, [2014] 1 SCR 704 41
Remuneration of Judges (No 2), In Re [1998] 1 SCR 4 (SCC) 111

Saltspring Island Water Preservation Society v Rockliffe 1993 CanLII 1313 (BC CA), [1993] 4 WWR 601 116
Schachter v Canada [1992] 2 SCR 679 (SCC) 102, 104, 105, 113, 117, 134, 169, 319
Severn v The Queen [1878] 2 SCR 70 (SCC) 36, 102
Sinclair v Quebec [1992] 1 SCR 579 (SCC) 111

Tolofson v Jensen 1994 CanLII 44 (SCC), [1994] 3 SCR 1022 332
Trociuk v British Columbia (Attorney General) [2003] 1 SCR 835 (SCC) 111
Turigan v Alberta 1988 ABCA 333, [1988] 6 WWR 673 109

UFCW v KMart Canada [1999] 2 SCR 1083 (SCC) 111

Valin v Langlois [1879] 3 SCR 1 (SCC) 36, 102
Vriend v Alberta [1998] 1 SCR 493 (SCC) 105

India

Ashoka Kumar Thakur v Union of India [2008] INSC 614 157

Basheshar Nath v Commissioner of Income Tax [1958] INSC 117, [1959] 1 Suppl SCR 528 152
Bawa Singh v Union of India 6 (1970) DLT 409 162
Behram Khurshed Pesikaka v State of Bombay [1954] INSC 15, [1955] 1 SCR 613 147–151, 160, 337
Bhikaji Narain Dhakras v State of Madhya Pradesh [1955] INSC 48, [1955] 2 SCR 589 ..... 151, 154, 158, 162, 245, 247

Chembakave Vadakkekkara Lakshmi v Nellisseri Gramam Narayanaswami AIR 1963 Ker 330 162
Chotey Lal v The State of Uttar Pradesh AIR 1951 All 228 46

Deep Chand v State of Uttar Pradesh [1959] INSC 3, [1959] Suppl SCR 8 ..... 152–154, 159, 162
Dharam Pal v Kaushalya Devi AIR 1990 Raj 135 162
Dwarkadas Shrinivas of Bombay v The Sholapur Spinning & Weaving Co Ltd [1953] INSC 87 674 150

Golaknath v State of Punjab [1967] INSC 45, [1967] 2 SCR 762 197

Hariharan v Reserve Bank of India [1999] INSC 37 144

Jagannath v Authorised Officer, Land Reforms [1971] INSC 279, [1972] 1 SCR 1055 ..... 154, 155

Keshavan Madhava Menon v State of Bombay [1951] INSC 3, [1951] SCR 228 ... 145, 146, 154, 200
KK Poonacha v State of Karnataka [2010] INSC 706 ..... 160, 162, 245

Mahendra Lal Jaini v State of Uttar Pradesh [1962] INSC 303, [1963] Suppl SCR 912 ..... 153, 154, 159, 162, 303
Minoo Framroze Balsara v Union of India AIR 1992 Bom 375 162

PL Mehara v DR Khanna AIR 1971 Delhi 1 162
Purshottam Govindji Halai v Shree BM Desai [1955] INSC 53, [1995] 2 SCR 887 . 158

R M D Chamarbaugwala v Union of India [1957] INSC 32, [1957] SCR 930 . 144, 145, 169
Rakesh Vij v Raminder Pal Singh Sethi [2005] INSC 522 ..... 161, 162, 245
Ram Chand v State of Haryana (1971) 73 PLR 958 162

Saghir Ahmad v State of UP [1954] INSC 89, [1955] SCR 707 . 153, 154, 172, 213, 245
State of Bombay v FN Balsara [1951] INSC 38, [1951] SCR 682 ..... 144, 147
State of Gujarat v Shri Ambica Mills [1974] INSC 72, [1974] 3 SCR 760 155–157, 162, 163, 249
State of Kerala v NM Thomas [1995] INSC 178, [1995] SCC 3 434 226
State of Kerala v NM Thomas [1975] INSC 224, [1976] SCR 1 906 144
State of Madhya Pradesh v Thakur Bharat Singh 1967 AIR 1170, [1967] 2 SCR 454 154
State of Rajasthan v Hindustan Sugar Mills Ltd [1988] INSC 174, [1988] 1 Suppl SCR 461 161
State of West Bengal vs Anwar All Sarkarhabib [1952] SCR 284 (INSC) 144
Sundararamier v State of Andhra Pradesh [1958] INSC 19, [1958] SCR 1422 151, 152, 161, 162
Sunil Batra v Delhi Administration [1978] INSC 148, [1979] SCR 1 392 144

The United States

Ackermann v United States 340 US 193 (1950) 62
Adkins v Children's Hospital 561 US 525 (1923) 95, 140, 213
Alaka Airlines Inc v Brock 480 US 678 (1987) 59
Allen v City of Louisiana 103 US 80 (1881) 59
Allison v Corker 67 NJL 596 (1902) 97
Almeida-Sanchez v United States 413 US 266 (1973) 71
American Trucking Associations v Scheiner 483 US 266 (1987) 79
American Trucking Associations v Smith 496 US 167 (1990) 79, 302
Apprendi v New Jersey 530 US 466 (2000) 84
Ayotte v Planned Parenthood of Northern New England 546 US 320 (2006) ... 91

Bacchus Imports v Dias 468 US 263 (1984) 79

Batson v Kentucky 476 US 79 (1986) 83
Bloom v Illinois 391 US 194 (1968) 71
Blount v Rizzi 400 US 410 (1971) 58
Blount v Rizzi 465 US 728 (1984) 58
Burgess v Salmon 97 US 381 (1878) 43
Byrd v Blue Ridge Rural Electric Cooperative 356 US 525 (1958) 76

Calder v Bull 3 US (3 Dall) 386 (1987) 47
Califano v Westcott 443 US 76 (1979) 58, 171
Champlin Refining Co v Corp Commission of Oklahoma 286 US 201 (1932) . 59
Chevron Oil Co v Huson 404 US 97 (1971) 76–79, 86
Chicago v Robbins 67 US 418 (1862) 63
Chicago, Indianapolis & Louisville Railway Company v Hackett 228 US 559 (1913) 94
Chicot County Drainage District v Baxter State Bank 308 US 371 (1940) .. 80, 81, 130
Cipriano v City of Houma 395 US 701 (1969) 80
Citizens United v Federal Election Commission 130 S Ct 876 (2010) 89, 94
City of Chicago v Morales 527 US 41 (1999) 88, 90
Commissioner v Wilcox 327 US 404 (1946) 64
Connolly v Union Sewer Pipe Co 184 US 540 (1902) 59

Davis v Michigan Department of Treasury 489 US 803 (1989) 86
Davis v Moore 772 A2d 204 (DC 2001) 82
Desist v United States 394 US 244 (1969) 72
DeStefano v Woods 392 US 631 (1968) 71
Dombrowski v Pfister 380 US 479 (1965) 49
Douglass v County of Pike 101 US 677 (1879) 76
Duncan v Louisiana 391 US 145 (1968) 71

Erie R Co v Tompkins 304 US 64 (1938) 64, 65, 70, 244
Escobedo v Illinois 378 US 478 (1964) 71
Eubanks v Wilkinson 937 F2d 1118 (6th Cir 1991) 49

Fay v Noia 372 US 391 (1963) 70
Frank v Mangum 237 US 309 (1915) 47
Free Enterprise Fund v Public Co Accounting Oversight Board 561 US 477 (2010) 59, 94
Frontiero v Richardson 411 US 677 (1973) 58
Fuller v Alaska 393 US 80 (1968) 72

Gardner v The Collector 73 US 499 (1868) 43
Gasperini v Center for Humanities 518 US 415 (1996) 76
Gelpcke v City of Dubuque 68 US (1 Wall) 175 (1863) 64, 68, 76
Gilbert v State of California 388 US 263 (1967) 70, 71
Gonzales v Carhart 127 S Ct 1610 (2007) 91, 92

Gonzales v Raich 545 US 1 (2005) 88
Goodyear Tire and Rubber Company v Vinson 749 So2d 393 (Ala 1999) 82
Great Northern Railway Co v Sunburst Oil 287 US 358 (1932) 64, 68
Greene v Fisher 132 S Ct 38 (2011) 85
Griffin v California 308 US 609 (1965) 71
Griffith v Kentucky 479 US 314 (1987) 50, 75, 78, 83–85, 252
Gulf Offshore Co v Mobil Oil Corp 453 US 473 (1981) 77

Hanna v Plumer 380 US 460 (1965) 76
Harper v Virginia Department of Taxation 509 US 86 (1993) .. 50, 76, 85, 86, 252
Harris v Arizona Independent Redistricting Commission 578 US — (2016) ... 94
Hepburn v Griswold 75 US 603 (1868) 96
Horn v Banks 536 US 266 (2002) 85

Iowa-Des Moines Nat Bank v Bennett 284 US 239 (1931) 171

James B Beam Distilling v Georgia 501 US 529 (1991) 79, 80, 86, 300, 337
James v United States 366 US 213 (1961) 64
Janklow v Planned Parenthood 517 US 1174 (1995) 90
Jawish v Morlet 86 A2d 96 (Colum App Ct 1952) 49
Jeter v Hewitt 63 US 352 (1859) 61
Johnson v New Jersey 384 US 719 (1966) 71

Katz v United States 389 US 347 (1967) 72
Kaufman v United States 394 US 217 (1969) 70
Kopp v Fair Political Practices Commission 11 Cal 4th 607 (1995) 49
Kuhn v Fairmont Coal Co 215 US 349 (1910) 64
Kunz v New York 340 US 290 (1951) 89

Landgraf v USI Film Producers 551 US 224 (1994) 49, 78, 86
Lanzetta v New Jersey 306 US 451 (1939) 89
Lee v Florida 392 US 378 (1968) 72
Legal Tender Cases 79 US 457 (1870) 96, 213
Leisy v Hardin 135 US 100 (1890) 95, 96, 213
Lemon v Kurtzman 411 US 192 (1973) 66
Linkletter v Walker 381 US 618 (1965) 65, 67–73, 75, 83, 84, 313
Loffland Bros Co v Roberts 386 F2d 540 (5th Cir 1968) 77
Lovell v City of Griffin 303 US 444 (1938) 89

Mackey v United States 401 US 667 (1971) 74, 75, 84, 132, 337
Mapp v Ohio 367 US 643 (1961) 68, 70, 71
Marbury v Madison 5 US 137 (1803) 47, 56, 62, 63, 80, 82, 94, 139, 244
Massachusetts v Mellon 262 US 447 (1923) 49, 150, 245, 250
Matthews v Zane 20 US 164 (1822) 43
Minnesota Voters Alliance v Mansky 585 US — (2018) 66
Miranda v Arizona 384 US 436 (1966) 71
Mosser v Darrow 341 US 267 (1951) 64

Movible Offshore Co v Ousley 346 F2d 870 (5th Cir 1965) 77
Murphy v National Collegiate Athletic Association 584 US — (2018) 49, 60–62, 66

Norton v Shelby County 118 US 425 (1886) 63, 80–82, 94, 244

Ohio Life Insurance and Trust Co v Debolt 57 US 416 (1853) 76

Penry v Lynaugh 492 US 302 (1989) 84
Perez v Ledesma 401 US 82 (1971) 49, 94
Pidgeon v Turner 538 SW3d 73 (Tex 2017) 49
Planned Parenthood of Southeastern Pennsylvania v Casey 505 US 833 (1992) 92, 303
Plessy v Ferguson 163 US 537 (1896) 298
Polites v United States 364 US 433 (1960) 62
Pure Oil Co v Snipes 57 F2d 416 (5th Cir 1961) 77

Rahrer, In Re 140 US 545 (1891) 95, 96, 213
Reno v Flores 507 US 292 (1993) 90
Reynoldsville Casket Co v Hyde 514 US 749 (1995) 86
Reynoldsville Casket Company v Hyde 514 US 749 (1995) 94
Ring v Arizona 536 US 584 (2002) 84
Rodrigue v Aetna Casualty & Surety Co 395 US 352 (1969) 77
Rowan v Runnels 46 US 134 (1847) 76

Sabri v United States 541 US 600 (2004) 90
Saint Francis College v Al-Khazraji 481 US 604 (1987) 77
Skinner v Oklahoma 316 US 535 (1942) 171
State Ex Rel Moore v Molpus 578 So2d 624 (Ala 1991) 82
State v Hodge 128 St3d 1 (Ohio 2010) 82
Steffel v Thompson 415 US 452 (1974) 49, 94
Stenberg v Carhart 530 US 914 (2000) 92
Stovall v Denno 388 US 293 (1967) 70–72, 83
Swift v Tyson 41 US 1 (1842) 63, 64

Teague v Lane 489 US 288 (1989) 50, 75, 84, 85, 252
Tehan v Shott 382 US 406 (1966) 71

United States v Delaware & Hudson Co 213 US 366 (1909) 57
United States v Johnson 457 US 537 (1982) 72, 75, 77
United States v Peltier 422 US 531 (1975) 71, 73
United States v Raines 362 US 17 (1960) 91
United States v Salerno 481 US 739 (1987) 89, 90, 92, 93, 250
United States v Stevens 559 US 460 (2010) 92
United States v The Schooner Peggy 5 US (1 Cranch) 103 (1801) 49, 77, 78
United States v Wade 388 US 218 (1967) 70, 71

Virginia v American Booksellers Association 484 US 383 (1988) 58

Washington v Glucksberg 521 US 702 (1997) 90
Weinberger v Wiesenfeld 420 US 636 (1975) 58
Welsh v United States 398 US 333 (1970) 171
West Coast Hotel Co v Parrish 300 US 379 (1937) 95, 140, 213
Whole Woman's Health v Hellerstedt 579 US — (2016) 303
Williams v United States 401 US 646 (1971) 72
Winsness v Yocom 433 F3d 727 (10th Cir 2006) 49

Yazoo & Mississippi Valley Railroad v Jackson Vinegar 226 US 217 (1912) 89

South Africa

C v Department of Health and Social Development, Gauteng [2012] ZACC 1, (2012) 2 SA 208 167
Camps Bay Ratepayers' Association v Harrison [2010] ZACC 26, (2011) 4 SA 113 176
Camps Bay Ratepayers' Association v Harrison [2010] ZACC 19, (2011) 4 SA 42 . 176
Case v Minister of Safety and Security; Curtis v Minister of Safety and Security [1996] ZACC 7, (1996) 3 SA 617 169
Coetzee v Government of the Republic of South Africa; Matiso v Commanding Officer Port Elizabeth Prison [1995] ZACC 7, (1995) 4 SA 631 169

Dawood v Minister of Home Affairs [2000] ZACC 8, (2000) 3 SA 936 33, 180–182
De Vos NO v Minister of Justice and Constitutional Development [2015] ZACC 21, (2015) 9 BCLR 1026 182
Department of Transport v Tasima (Pty) Ltd [2016] ZACC 39, (2017) 1 BCLR 1 . 174–177, 244
Doctors for Life International v Speaker of the National Assembly [2006] ZACC 11, (2006) 6 SA 416 180

Economic Freedom Fighters v Speaker of the National Assembly [2016] ZACC 11, (2016) 3 SA 580 177
Executive Council of Western Cape Legislature v President of the Republic of South Africa [1995] ZACC 8, (1995) 4 SA 877 180

Ferreira v Levin [1995] ZACC 13, (1996) 1 SA 984 ..... 172, 173, 243, 244
Fose v Minister of Safety and Security [1997] ZACC 6, (1997) 3 SA 786 172
Fourie v Minister of Home Affairs [2004] ZASCA 132, (2005) 1 All SA 273 .. 197
Fraser v Children's Court Pretoria North [1997] ZACC 1, (1997) 2 SA 218 ... 180

Gaertner v Minister of Finance [2013] ZACC 38, (2014) 1 SA 442 168
Geldenhuys v State [2008] ZASCA 47, (2009) 1 LRC 294 ..... 173, 243
Gory v Kolver NO [2006] ZACC 20, (2007) 4 SA 97 173

Ingledew v Financial Services Board [2003] ZACC 8, (2003) 4 SA 584 173

Islamic Unity Convention v Independent Broadcasting Authority [2002] ZACC 3, (2002) 4 SA 294 170

J v Director General, Department of Home Affairs [2003] ZACC 3, (2003) 5 SA 621 ..... 180, 182
Jaftha v Schoeman; Van Rooyen v Stoltz [2004] ZACC 25, (2005) 2 SA 140 .. 168
Janse van Rensburg v Minister of Trade and Industry [2000] ZACC 18, (2001) 1 SA 29 181

Khosa v Minister of Social Development; Mahlaule v Minister of Social Development [2004] ZACC 11, (2004) 6 SA 505 167
Kouga Municipality v Bellingan [2011] ZASCA 222, (2012) 2 SA 95 178
Kruger v President of the Republic of South Africa [2008] ZACC 17, (2009) 1 SA 417 171

Lawyers for Human Rights v Minister of Home Affairs [2017] ZACC 22, (2017) 5 SA 480 181
Levenstein v Estate of the Late Sidney Lewis Frankel [2018] ZACC 16 ... 239, 240

MEC for Health, Eastern Cape v Kirland Investments (Pty) Ltd [2014] ZACC 6, (2014) 3 SA 481 ..... 174, 176
Merafong City Local Municipality v AngloGold Ashanti Ltd [2016] ZACC 35 176
Minister of Communications v Ngewu [2013] ZACC 444, (2014) 3 BCLR 364 315
Moise v Greater Germiston Transitional Local Council [2001] ZACC 2, (2001) 4 SA 1288 ..... 173, 238

National Coalition for Gay and Lesbian Equality v Minister of Home Affairs [1999] ZACC 17, (2000) 2 SA 1 ..... 171, 303
National Coalition for Gay and Lesbian Equality v Minister of Home Affairs [1999] ZACC 17, (2000) 2 SA 1 ..... 166, 167

Oriani-Ambrosini, MP v Sisulu, MP Speaker of the National Assembly [2012] ZACC 27, (2012) 6 SA 588 169
Oudekraal Estates (Pty) Ltd v City of Cape Town [2004] ZASCA 48, (2004) 3 All SA 1 ..... 174, 176, 337

Prince v President, Cape Law Society [2000] ZACC 28, (2001) 2 SA 388 173
Provincial Minister for Local Government, Environmental Affairs and Development Planning, Western Cape v Municipal Council of the Oudtshoorn Municipality [2015] ZACC 24, (2015) 6 SA 115 168

Ramuhovhi v President of the Republic of South Africa [2017] ZACC 41, (2018) 2 SA 1 166

S v Bhulwana; S v Gwadiso [1995] ZACC 11, (1996) 1 SA 388 179
S v Steyn [2000] ZACC 24, (2001) 1 SA 1146 181
Sarrahwitz v Martiz NO [2015] ZACC 14, (2015) 4 SA 491 169

Shoprite Checkers (Pty) Limited v Member of the Executive Council for Economic Development [2015] ZACC 23, (2015) 6 SA 125 173
South African National Defence Union v Minister of Defence [2007] ZACC 10, (2007) 5 SA 400 180
South African Transport and Allied Workers Union v Garvas [2012] ZACC 13, (2013) 1 SA 83 169

Teddy Bear Clinic for Abused Children v Minister of Justice and Constitutional Development [2013] ZACC 35, (2014) 2 SA 168 ..... 169, 182

Zondi v MEC for Traditional and Local Government Affairs [2004] ZACC 19, (2005) 3 SA 589 181

Other Jurisdictions

Amministrazione dello Stato delle Finanze v Simmenthal SpA (106/77) [1978] ECR 629 345

Table of Legislation

Ireland

Criminal Law (Amendment) Act 1935
s 1
(1) 128

Finance Act 1980 125

Income Tax Act 1967 125

Misuse of Drugs Act 1977
s 2
(2) 134

Offences Against the State Act 1939
s 29
(1) 131
s 34 126

Statute of Limitations 1957
s 11
(2) 126

England and Wales

Colonial Laws Validity Act 1965
s 2 36

Statute of Westminster 1931 36

South Africa

Road Accident Fund Amendment Act 2005
s 4 171
s 6 171
s 10 171
s 11 171
s 12 171

Table of Constitutional Provisions

Ireland

Constitution of Ireland 1937
- art 15.4 ..... 30, 145
- art 13
  - (3) 44
- art 15 ..... 122, 124, 127, 243
  - (4) 31, 123, 136
- art 24 122
- art 25 ..... 43, 124, 127
  - (4) 123
- art 26 ..... 40, 44
  - (1) 39
  - (2) 39, 40
  - (3) 39
- art 34 140
  - (3) 31, 39
- art 50 ..... 30, 31, 38, 136

Canada

Canadian Constitution Act 1982
- s 24
  - (1) 29
- s 52 248
  - (1) 29, 38, 111, 145

Constitution Act 1867
- s 55 44
- s 56 44
- s 57 44
- s 91 36
- s 92 36
- s 133 111

Manitoba Act 1870
- s 23 111

United States

Constitution of the United States 1787
(1)(9)(3) 47
art 1
(7) 43
(8) 34
art 3 72
(2) 75

India

Constitution of India 1950
art 13 ..... 28, 38, 145, 162
(1) 149, 150, 161
(2) 149
art 32 28
art 111 45
art 131 29
art 132
(1) 29
art 133
(1) 29
art 134 29
art 137 149
art 143 ..... 42, 46
art 245 36
art 246 36

European Union

Treaty on the Functioning of the European Union (as amended) 1957
art 267 345

South Africa

Constitution of South Africa 1996
s 2 174
s 79 46
s 81 46
s 149 ..... 33, 34, 37
s 167
(4) 32
s 168
(3) 32

s 172 134
(1) 32–34, 174, 178, 179
s 192 170

Other Constitutions

Constitution of South Africa 1996
s 167
(4) 86
s 168
(3) 87
s 172
(1) 86, 87

List of Figures

Figure 2.1 Overview of US Law on Retroactivity 50

Figure 9.1 The Inclusion Model 268
Figure 9.2 The Membership Model 268
Figure 9.3 Model of a Legal System 279

Figure 10.1 The Retrocausality Problem: Invalidity and Changing Law . 306
Figure 10.2 Superposition Model of Legal Validity 311

List of Tables

Table 8.1 Combinations of Answers to Questions on Unconstitutionality (Strikethrough Indicates Views Containing Tensions) 241

Part I.

Introduction

1 Introduction

CHAPTER OVERVIEW

1.1. Subject of the Thesis 4
1.1.1. The Problem of Unconstitutionality 6
1.2. Methodology and Limitations 7
1.2.1. Functionalism and Universalism 8
1.2.2. Situating the Thesis in Comparative Scholarship 9
1.2.3. Jurisdiction Selection 11
1.2.3.1. Strong and Weak Judicial Review 12
1.2.3.2. Centralised and Decentralised Courts 13
1.2.3.3. Concrete and Abstract Judicial Review 15
1.2.3.4. Jurisdictions Selected for Study & Limitations of the Thesis 16
1.2.3.5. Geographic Dispersal and Grounding of Judicial Review 18
1.2.4. The Role of Legal Theory 19
1.3. Structure of the Thesis 20
1.3.1. Part One: Introduction 20
1.3.2. Part Two: Declarations of Unconstitutionality in Practice . 20
1.3.3. Part Three: Theoretical Analysis 21
1.3.4. Part Four: Conclusion 23

1.1. SUBJECT OF THE THESIS

This thesis is a comparative and theoretical study of declarations of unconstitutionality in the context of judicial review of legislation.[1] It seeks to answer two central questions:

(1) What is the nature of unconstitutionality being declared, and what does this declaration achieve?
(2) What ought the effect of a declaration of unconstitutionality be?

The first question is answered through a descriptive and comparative analysis that assists theory-formation. The data generated by this analysis are used to ascertain to what extent, if any, there are significant points of convergence or divergence in the practice of unconstitutionality in legal systems sharing a common law tradition. The comparisons and contrasts in these practices are used to generate theoretical models that describe how unconstitutionality manifests in more nuanced terms. This synthesis draws attention to key details and trends, arguing that, while unconstitutionality may have certain general and transnational features, it also necessarily involves more local and idiosyncratic choices. The second, normative question is then answered by criticising the models on practical and theoretical grounds, and by advocating for a new model of unconstitutionality that overcomes these criticisms. Both questions are asked against the background of certain methodological constraints, which are explained in greater detail in this chapter. The most considerable of these constraints is that of jurisdiction selection; specifically, this thesis considers unconstitutionality as it operates in jurisdictions that have a common law tradition and that practice strong-form judicial review.

This is not the first comparative study addressing unconstitutionality in such jurisdictions; comparison between Ireland, the USA, and Canada has been made before.[2] Outside these methodological constraints, comparison has also been drawn

[1] Where this thesis speaks of 'judicial review' without qualification, it should be assumed that it refers to judicial review of legislation.

[2] William Mark Murphy, 'The Problem of Unconstitutionality and Retroactivity in Criminal Law: Ireland, the US and Canada Compared' (2007) 42 Ir Jur 63.

between West Germany and Canada.[3] Attempts to systematically theorise the practice of single jurisdictions have also been undertaken.[4] However, the issue of unconstitutionality is often wholesale omitted from modern comparative constitutional law handbooks, except insofar as it relates to another substantive topic.[5] Moreover, close scholarly attention has not been paid to the stark theoretical differences that occur between jurisdictions. This has left unconstitutionality somewhat insufficiently theorised. This is also partly a feature of the relative dearth of wide-ranging comparative study. Even in the common law world, prior studies have not tended to look further afield at jurisdictions such as India and South Africa. Recent comparative scholarship in this field has tended to focus more on the political context of judicial review than the technical effects of unconstitutionality.[6] The project is therefore novel in the breadth of jurisdictions covered and its particular focus on unconstitutionality.

The thesis also presents a novel framework for unconstitutionality. The comparative data are combined with often-overlooked theoretical points about the nature of legal systems to develop a sophisticated analysis of unconstitutionality. The models that emerge from this analysis represent a new way of understanding the phenomenon of unconstitutionality, as well as mounting more nuanced criticism of practice. This framework thus enables greater theoretical insight to this area of law, and would be a useful point of departure for future work that seeks to generalise the findings of this thesis outside its methodological restrictions.

---

[3] Susan Gluck Mezey, 'Civil Law and Common Law Traditions: Judicial Review and Legislative Supremacy in West Germany and Canada' [1983] International and Comparative Law Quarterly 689.

[4] Oliver P Field, *The Effect of an Unconstitutional Statute* (University of Minnesota Press 1935).

[5] Michael Rosenfeld and Andras Sajo (eds), *The Oxford Handbook of Comparative Constitutional Law* (OUP 2012); Tom Ginsburg and Rosalind Dixon (eds), *Comparative Constitutional Law* (Edward Elgar 2013).

[6] See for example: Erin Delaney and Rosalind Dixon (eds), *Comparative Judicial Review* (Edward Elgar 2018) (though Virgílio Afonso da Silva's contribution does make some note of the divergent effects achieved by judicial review in some jurisdictions).

1.1.1. The Problem of Unconstitutionality

The issues surrounding unconstitutionality become particularly acute when considered through the lens of a case study. Consider two jurisdictions/legal systems operating, at different points in time, in the same geographical region 'the Island of Ireland': the Irish Free State and Ireland.[7] In the year 1935, the Irish Free State has legislative authority on the Island of Ireland, and it enacts a law providing for a criminal offence. Two years later, in 1937, Ireland declares its independence, breaking away from the Irish Free State. In the course of declaring its independence, however, Ireland provides that it will presumptively, subject to their conformity with the norms in its new Constitution, carry over all the laws of the Irish Free State that were effective on the Island of Ireland (which territory is unchanged through this legal revolution).

The basis for the continuing legal authority of the Irish Free State's 1935 statute on the Island of Ireland is now Ireland's 1937 Constitution. This arrangement continues for almost 100 years, until 2006 when the Supreme Court of Ireland finds the 1935 statute unconstitutional. This gives rise to a difficult question of how Ireland must now characterise this law. The 1935 statute provides for a criminal offence that has been applied in Ireland for three quarters of a century. Yet, Ireland's Supreme Court has now declared that the offence suffers from a serious constitutional defect. Because the statute was only *presumed* to be constitutional in 1937, the Court must declare that this presumption has been rebutted, and therefore that the 1935 law was not validly carried over by Ireland's 1937 Constitution. Does this mean that, from 1937, the offence detailed in the 1935 statute was, in fact, not an offence in the state of Ireland? Where does this leave citizens of Ireland who were punished on foot of this law?

This example is not simply an abstraction. It details a real and pressing legal issue

[7] This is the constitutional name for the Irish state. The name 'The Republic of Ireland' only came later with the passage of the Republic of Ireland Act 1948, and it is not a constitutional title for the state.

that confronted the Supreme Court of Ireland in 2006.[8] Similar alarming issues have been raised in other apex constitutional courts; the Supreme Court of Canada was asked, early after its establishment, whether the vast majority of the Manitoba statute book was null and void because it had not been properly translated into French.[9] These cases cast in sharp relief the question of how we characterise unconstitutional laws.

1.2. METHODOLOGY AND LIMITATIONS

This thesis is not intended to be a large-data empirical study. Such an undertaking would be useful in generating more stark comparison and contrast across a wider number of variables and jurisdictions. However, it would risk not doing justice to some subtleties in the operation of unconstitutionality that are often missed. It would also likely be too ambitious an undertaking within the constraints of a single doctoral thesis. This requires that there be some limitations on candidates for analysis.

Some difficulties that present themselves in selecting jurisdictions are practical ones, such as language barriers and the availability of legal materials.[10] Others are principled limitations. For one, I suggest that case law in judicial review is more likely to shed light on the conceptual foundations of unconstitutionality if judges are required to frame their decisions against some sense of what law is, rather than being able to present their decisions as more openly political choices. This is more likely in countries practising decentralised, strong-form judicial review. I explore the grounds for these criteria further below. Additionally, keeping institutional design choices relatively consistent across the systems to allow for more provocative com-

[8] *CC v Ireland* [2006] IESC 33, [2006] 4 IR 1.

[9] *In Re Manitoba Language Rights* [1985] 1 SCR 721 (SCC).

[10] As Jackson has conceded: 'comparator countries to be studied may be limited by the languages the scholar is familiar with, or the accessibility of the legal information.' Vicki Jackson, 'Comparative Constitutional Law: Methodologies' in Michael Rosenfeld and Andras Sajo (eds), *The Oxford Handbook of Comparative Constitutional Law* (OUP 2012) 64.

parison as it makes contrasts and points of divergence even more stark and theoretically interesting.

This section first outlines my approach to comparative methodology in an abstract sense, situating it both on the 'functionalist'/'universalist' spectrum and in a framework recently proposed by Leckey for categorising comparative scholarship. After this, it describes my criteria for jurisdiction selection. The section then concludes with some brief remarks on the role to be played by legal theory in this thesis, and how such theoretical reflection interacts with the data generated by the comparative study.

1.2.1. Functionalism and Universalism

Comparative constitutional scholarship can be divided into two types of approach: a 'functionalist' approach, which looks at systems and institutions, and a 'universalist' approach, which attempts to derive certain truths about some aspect of law (human rights is a particularly good example of this) that should prevail in every legal system.[11] The approach of this thesis is not easy to classify sharply under this binary distinction. However, it is fair to say that it leans more towards a normatively-inclined functionalist analysis than a universalist one. Consider the following quote from Ackerman, which Jackson holds up as an exemplar of the functionalist approach:

> My aim is to identify (a) one or another common problem confronting different 'constitutional courts,' and then follow up by specifying (b) different coping strategies these courts have adopted as they have tried to solve the problems. Once we have gained some clarity on these two issues, we may hope for a deeper insight into the comparative value of competing coping strategies.
>
> … Much of the best comparative scholarship follows a similar method, first defining a common problem—for example, the protection of freedom of speech—and then considering different doctrinal solutions proposed by different courts, before passing a considered judgment on the best approaches.[12]

[11] Jackson (n 10) 61–66.

[12] Bruce Ackerman, 'The Rise of World Constitutionalism' (1997) 83 Va L Rev 771, 794. Cited in: Jackson (n 10) 63.

My approach for much of this thesis reflects the same goal and approach. I am concerned primarily with identifying and evaluating various coping strategies that address difficulties arising from the practice of judicial review of legislation. I am, however, also concerned with analysing this practice in such a way as to shed light on fundamental jurisprudential questions that are germane to unconstitutionality. In particular, questions around the nature of law can partly be illuminated through a comparative study of the life and death of law, which is something unconstitutionality must incorporate.

There is, in summary, both a functionalist and universalist edge to the comparative methodology I employ here. I not only seek to uncover an answer to 'best practice' from a functional point of view (what works best within a sound constitutional framework) but also from a universalist point of view (what theory of unconstitutionality best fits with a sound theory as to the nature of law generally).

1.2.2. Situating the Thesis in Comparative Scholarship

In a recent review article, Leckey has set out at least thirteen different modes of comparative law.[13] These include the following, though for the sake of brevity I do not list all thirteen:

> (6) As a scholar, studying national laws with a view to identifying their common core.
>
> [...]
>
> (11) As a scholar, studying law from more than one jurisdiction—often including the author's own—with the aim of 'better knowledge of legal rules and institutions', and perhaps of law's surrounding society.
>
> (12) As a scholar, studying together the law of two or more jurisdictions, with the aim of better understanding law generally as an intellectual practice or as a form of social ordering underwritten by claims of right and legitimacy.

[13] Robert Leckey, 'Review of Comparative Law' (2017) 26 Social & Legal Studies 3.

This is by no means the only taxonomy of comparative approaches.[14] It does, however, have the advantage of being very recent in a field that has grown and developed significantly in the last decade or two in an increasingly globalised world. For that reason, it is useful to situate my approach within this framework.

This thesis fits somewhere in approaches (6), and either (11) or (12). The main difference between these approaches for present purposes is that (6) often implies a reformist bent, whereas (11) and (12) are descriptive and aimed at deepening understanding. Much of the thesis could be read in this relatively non-committal way. Only towards the conclusion does it engage in an exercise more properly captured by sense (6). This approach can be criticised on the basis that transferring legal approaches outside of their jurisdictions is illegitimate at worst, and difficult at best.[15]

I would suggest that the transplantation of approaches is more legitimate here than perhaps in other areas of public law. In many cases, constitutional disputes are fraught with political and moral challenges. Questions of substantive rights most obviously fit this mould. Unconstitutionality, by contrast, is a more technical exercise. To be sure, the *consequences* of unconstitutionality pose political challenges, but these challenges are not particularly idiosyncratic. The chief difficulty is nearly always the risk of a gaping lacuna in the law. Some laws will leave more significant lacunae than others, but there is nothing particularly unique or special about filling gaps as a technical exercise. Thus, this thesis does not require the imposition of one moral or political outlook. It takes as granted some politically contentious ideas, such as strong-form judicial review; however, once these assumptions are granted, unconstitutionality becomes more technical. Moreover, the jurisdictions selected share a common law heritage, and theoretical perspectives on the nature of law in these systems have been predominantly shaped by Anglophone scholars in the legal positivist tradition. As unconstitutionality is intimately bound up with questions

[14] For example, see also: Günter Frankenberg, 'Critical Comparisons: Re-thinking Comparative Law' (1985) 26 Harv Int'l L J 411; Günter Frankenberg, 'The Innocence of Method – Unveiled: Comparison as an Ethical and Political Act' (2014) 9 J Comp L 222.

[15] Leckey, 'Review of Comparative Law' (n 13) 9–10.

around the nature of law, the significance of this unity should not be underestimated.

1.2.3. Jurisdiction Selection

As it is, in part, a study of judicial review, one of the principal factors to be isolated in this thesis is the role played by the judicial power. Because this requires a study of judicial practice, it is important to recognise the distinctions in the judicial role between civil and common law judges.[16] Because the conceptions of the role of the judge are so different within these traditions, comparison between them can be complex and difficult. This thesis focuses on the common law tradition because of the general familiarity of common law judges with law-making through the interpretation of legally authoritative materials. This interpretative attitude is reflected to a greater extent in the judgments of common law judges than their civilian counterparts. It also animates the difficulties with retrospective effect that can occur in constitutional cases where judges develop new constitutional rules. Future work might build on whether the models and theory proposed by this thesis apply equally to the civil law tradition, but that is outside of the scope of this thesis.

Constants for selected legal systems include: an established practice of strong-form judicial review of legislation,[17] concrete review of legislation rather than abstract review, and decentralised constitutional review rather than centralised constitutional review. These criteria are important for declarations of unconstitutionality for at least three reasons. Strong-form judicial review heightens the consequences of unconstitutionality, enabling the judiciary to 'strike down' unconstitutional law of its own motion. Concrete review of legislation entails unconstitutionality arising

[16] Civil law judges often hone more technical skills than their common law counterparts, and in some civil law jurisdictions (for example, France) judge and lawyer are even separate career tracks. Vicki Jackson and Mark Tushnet, *Comparative Constitutional Law* (2nd edn, Foundation Press 2006) 467.

[17] That is, a practice of judicial review that allows the judiciary to alter a property or properties of a piece of legislation without further recourse to another institution. I explain this term, and the corresponding 'weak-form' judicial review, in further detail below with reference to Tushnet's work.

in actual cases, and against the background of legislation with a history of being applied. This increases the ripple effects of unconstitutionality, rather than allowing unconstitutionality to be apprehended in the abstract before the legislation begins to have real-world effect. Finally, decentralised review entails that more courts hear constitutional matters. This gives rise to an increased possibility of court disagreements, and therefore revision of constitutional standards. If more courts hear disputes implicating the constitutionality of legislation it also means, more simply, that there will be more findings of unconstitutionality. This likely makes unconstitutionality more common, and thus more pressing, in decentralised systems.

1.2.3.1. Strong and Weak Judicial Review

In more recent times, a model of 'dialogic'[18] or 'weak-form'[19] judicial review has emerged as a challenger to traditional 'strong-form' judicial review. Tushnet outlines the shared traits of strong- and weak-form judicial review as follows:

> (1) The legislature enacts a statute;
> (2) The statute is challenged before the constitutional court;
> (3) The constitutional court holds the statute unconstitutional.[20]

So far there is common ground. The difference between the two forms of review is that in dialogic/weak-form review, the legislature is given an opportunity to respond to the finding in (3) and re-enact the statute (potentially according to some modified procedure). In strong-form judicial review, the only possible further fourth step is that a constitutional amendment reverses the court's finding.[21]

---

[18] Peter Hogg and Allison Bushell, 'The Charter Dialogue between Courts and Legislatures (Or Perhaps the Charter of Rights Isn't Such a Bad Thing after All)' (1997) 35 Osgoode Hall L J 75; Peter Hogg, Allison Bushell Thornton and Wade Wright, 'Charter Dialogue Revisited: Or "Much Ado About Metaphors"' (2007) 45 Osgoode Hall L J 1.

[19] Mark Tushnet, *Weak Courts, Strong Rights: Judicial Review and Social Welfare Rights in Comparative Constitutional Law* (Princeton University Press 2008).

[20] Mark Tushnet, *Advanced Introduction to Comparative Constitutional Law* (Edward Elgar 2014) 57.

[21] ibid 57. Even then, in certain jurisdictions it is not clear that this would automatically revive the statute. It may be the case that the legislature must re-enact the instrument, but it would only be free to do so after the amendment.

I am interested in strong-form review precisely because of its approach to the hypothetical fourth step here. Its hard-nosed approach to reversing the finding of the court, and thus the exceptional weight that strong-form review attaches to unconstitutionality, has caused issues in jurisdictions where it has been accepted. Indeed, it is this absolutist feature of strong-form judicial review—and the tension it can sow between the judiciary and legislature—that is meant to show how weak-form review is more attractive.[22] The consequences of unconstitutionality being particularly pronounced in jurisdictions that favour strong-form review makes them excellent candidates to explore both what these consequences are, and how they can potentially be mollified.

1.2.3.2. Centralised and Decentralised Courts

There are two primary models of constitutional court: the centralised and decentralised models.[23] Centralised systems are those in which only one organ—a bespoke constitutional court—has the power to strike down laws as unconstitutional. Decentralised systems are those in which several courts have jurisdiction to rule on the constitutionality of laws. In this study, I focus chiefly on decentralised systems, as such systems are generally less open to embracing the political significance of judges reviewing legislation. Appointment processes for centralised constitutional courts are often more expressly political than those that govern appointment to general, decentralised courts.[24] The problems of unconstitutionality are thus more likely to be cast in a formal legal mould in decentralised jurisdictions, to help preserve this

[22] Mark Tushnet, 'The Rise of Weak-Form Judicial Review' in Tom Ginsburg and Rosalind Dixon (eds), *Comparative Constitutional Law* (Edward Elgar 2011) 322–23.

[23] Mauro Cappelletti, 'Judicial Review in Comparative Perspective' (1970) 58 Ca L Rev 1017; Jackson and Tushnet (n 16) 465.

[24] 'In Europe, the selection processes for judges of constitutional courts are typically different from, and more political than, those used to choose ordinary judges.' Victor Ferreres Comella, 'The European model of constitutional review of legislation: Toward decentralization?' (2004) 2 International Journal of Constitutional Law 461, 469.

air of political detachment.[25]

Centralised systems also often favour abstract review. This is explained further below, but it is effectively review of legislation that occurs other than against the background of a specific case or set of facts. One consequence of this difference is that while centralised constitutional courts will often have specific rules regulating the temporal effects of their judgments, thus providing a straightforward legal answer to judicial discretion in mollifying the retrospectivity of declarations of unconstitutionality,[26] decentralised systems do not tend to have such rules in place (at least not in codified form). This requires individual judges in each jurisdiction to determine a suitably flexible rule for the validity, or invalidity, of unconstitutional legal norms.

Finally, decentralised review presents challenges that are independently interesting, and which sharpen the consequences of unconstitutionality when it arises. Decentralised review increases the possibility of unconstitutionality arising in the first place; more courts will hear constitutional issues and decide on them in a decentralised regime. Additionally, because a decentralised system entails a hierarchy of courts with constitutional jurisdiction it also animates the challenging issue of the distinct roles higher and lower courts have in shaping constitutional law and

---

[25] It is precisely this recognition of the political elements of judicial review that makes centralised systems wary of the power to strike down laws in the first place: 'Centralized judicial review is closely associated, philosophically, with notions of parliamentary supremacy and a corresponding suspicion of permitting judges to set aside laws.' Jackson and Tushnet (n 16) 466.
However, a centralised system could have a positive disposition towards judicial review and would thus generate comparatively more findings of unconstitutionality. The frequency of judicial review is, of course, partly dependent on legal and judicial culture. Japan is a good example of a centralised system with a marked reluctance to exercise a judicial review power. Law has observed that the *Saikō Saibansho* (最高裁判所) has been markedly reluctant to find legislation unconstitutional: David Law, 'The Anatomy of a Conservative Court: Judicial Review in Japan' (2008) 87 Tex L Rev 1545; David Law, 'Why has Judicial Review Failed in Japan' (2010) 88 Wash U L Rev 1425.

[26] This is where those courts allow review of enacted legislation at all. Some systems, such as France, historically did not: Gustavo Fernandes de Andrade, 'Comparative Constitutional Law: Judicial Review' (2001) 3 U Pa J Const L 977, 982. This changed only relatively recently; the Constitutional law on the Modernisation of the Institutions of the Fifth Republic, passed by the Parliament of France in July 2008, amended Article 61 of the French Constitution to allow for review of legislation post-enactment.

the possibility of disagreement between hierarchically-arranged courts.

1.2.3.3. Concrete and Abstract Judicial Review

There are two primary models of judicial review: concrete review and abstract review. Concrete review examines legislation for conformity with the constitution based on real cases taken by affected litigants. Abstract review usually takes place either through a constitutional challenge or a referral of a constitutional question. Constitutional challenges can be taken by various public institutions, such as the government, an ombudsman, or parliament.[27] As the review is abstract, these challenges do not require any proven hardship to have actually arisen. To counterbalance this, they usually must be taken within relatively strict time limits in countries that prefer this type of review. Constitutional questions, then, arise when a judge sitting in a court of inferior jurisdiction believes there may be an issue with the validity of a legal provision relevant to the case and states a question to be answered by the constitutional court.[28] The constitutional court in this case must still confine its analysis to the constitutionality of the queried legal provision, and it does not decide the entire case.

Regardless of how abstract review arises, a constant feature is that the court is never required to consider the impugned legal provision(s) as they apply to a particular case. This means that the potential fallout resulting from the law being unconstitutional is limited. Abstract review interferes with fewer, if any, vested rights. This makes retrospectivity a comparatively more pressing concern in concrete review cases. While the underlying theory of unconstitutionality should not differ materially across these cases, the practical consequences of unconstitutionality will. Because of this pressing need to avoid administratively catastrophic consequences of legal lacunae, jurisdictions that favour concrete review cast light on the practical problems that can arise from strong-form judicial review.

---

[27] Comella (n 24) 464.

[28] ibid 465.

1.2.3.4. Jurisdictions Selected for Study & Limitations of the Thesis

In the foregoing sections I have presented several constant factors that the selected legal systems must satisfy: they must be partly or completely rooted in the common law, they must allow for strong-form judicial review, they must prefer concrete to abstract judicial review, and they should have a decentralised system of judicial review. For the purposes of this study, therefore, the jurisdictions that I have selected are:

- Ireland: *common law, decentralised, strong-form concrete review*.
- India: *common law, decentralised, strong-form concrete review*.
- The United States: *common law, decentralised, strong-form concrete review*.
- Canada: *common law, decentralised, strong-form concrete review*.[29]
- South Africa: *mixed common and civil law, centralised, strong-form concrete review*.[30]

As I acknowledged above, there are limitations to a study such as mine. The focus is attenuated on the common law tradition. My study attempts to show how there is less convergence between these countries than might first be thought. However, an alternative comparative approach might attempt to start from a point of

[29] Classifying Canadian judicial review as 'strong-form' may seem contentious given the existence of the 'notwithstanding clause' in section 33. This allows individual provinces and territories to disapply the Charter with respect to their legislative enactments, if so they choose. Despite the intention behind the clause, it has only ever been exercised by two provinces (Quebec and Saskatchewan) and one territory (Yukon) and it has been argued that it has fallen entirely into desuetude: Richard Albert, 'Advisory Review: The Reincarnation of the Notwithstanding Clause' [2008] Alberta L Rev 1037; Aileen Kavanagh, 'A Hard Look at the Last Word' (2015) 35 OJLS 825.
Given the relative lack of impact achieved by the notwithstanding clause, it seems defensible to classify Canada's judicial review practice as strong-form in practice. Indeed, it is precisely because of this strong-form baseline that the Court pioneered the suspension of declarations of unconstitutionality.

[30] South Africa is chosen as it breaks slightly from the institutional framework articulated above; this is in the hope that it might provide interesting contrast while still broadly remaining within the methodological restrictions of the thesis.

greater difference (common law and civil law) or take a wholly different approach to classification of legal traditions within such families (for example, geography: one could compare Asian judicial review with Latin American judicial review). There is without a doubt interesting work to be done through such studies; however, due to limitations of both scope and resources I do not draw so wide a net here. The conclusions that emerge from this study must, therefore, be taken with a hint of caution; the breadth and generality of the claims made here may require more data and further support if applied outside of the context of the jurisdictions specifically analysed here. However, even within the scope of these limitations there are valuable lessons to be drawn from this thesis both for the comparative lawyer and for the legal theorist.

From the comparative perspective, it is of interest to identify the degree of homogeneity in practice among jurisdictions that are, at least formally and structurally, similar to one another. If there is a low degree of uniformity among the jurisdictions notwithstanding these similarities, it suggests that conscious design features of constitutional systems may not be effective in controlling the practice of unconstitutionality. This finding would have further consequences for the use of comparative law by constitutional courts; if there is more divergence in practices of unconstitutionality, this may have consequences for the utility of foreign law in deciding domestic issues around the operation of unconstitutionality.

From a theoretical perspective, a study of one of the ways in which a law can expire is instructive in understanding how laws exist generally. Legal philosophy is too often done in a vacuum, without the use of extensive data from various jurisdictions to support its claims. Often, legal philosophers will reference aspects of legal practice only in broad and abstract terms. This thesis aims to bring a greater level of doctrinal detail to this philosophical question, and buttress theoretical rigour with closer attention to the nuances and idiosyncrasies of legal practice in the surveyed jurisdictions. In sum, the thesis aims to advance theoretical understanding of declarations of unconstitutionality, at least as they are used in the surveyed jurisdictions. To be sure, this sample size is too small to be the last word on the issue, and future

work may explore the extent to which the arguments in this thesis generalise outside of its methodological limits.

1.2.3.5. Geographic Dispersal and Grounding of Judicial Review

Constitutional arrangements as well as geographical variety become the next most pressing concerns. Each of the jurisdictions chosen has a written constitution. Ireland and Canada form interesting comparators because they both have Bills of Rights, and they both provide for strong-form judicial review in their written constitutions. The US is like Canada in that it is a federal legal system with strong-form judicial review, but it lacks a textual basis for this review. This question of basis is important as constitutional texts often, at least notionally, resolve the question of the effect of unconstitutionality. The United States federal courts instead had to grapple with this question themselves, occasionally oscillating between two possibilities: erasing the legislation (voiding it), and disapplying it.

Similar to these jurisdictions, South Africa provides for strong-form judicial review in its Constitution. Unlike the others, however, it has a textual basis for remedies such as the suspended declaration of invalidity (which the Canadian courts developed judicially). In a sense, therefore, one might expect that Ireland, Canada, and South Africa mark out a spectrum proceeding from a less simple towards a more nuanced remedial response to judicial review of legislation. The textual basis in the South African case may also make such remedies relatively uncontroversial.

India is a close neighbour to Ireland regarding constitutional arrangements and secession from British colonial rule. Both Constitutions are reasonably old and both emerged from fraught colonial conditions. Both jurisdictions enable judicial review of legislation by a network of provisions that guarantee: judicial protection of fundamental rights, the voidness of laws inconsistent with either fundamental rights or federal laws (the latter only in India's case) and a subjection of Parliament to the provisions of the Constitution.

One would expect, then, that the approaches in Ireland and India would be broadly similar. In some respects, this is borne out (their approaches to the question

of review of pre- and post-Constitution laws are quite convergent, for example). In other respects, however, there is more marked divergence. Whereas India has followed the lead of the United States, Ireland has developed its approach with less conscious inspiration from other jurisdictions and arrived at a position more akin to that which operates in Canada. As shall be seen later, this marks a more profound divide than it might at first seem, and this thesis seeks to shed light on why this divergence occurs, taking account of comparative and theoretical data.

1.2.4. The Role of Legal Theory

My research seeks to apply a legal-theoretical framework of understanding to declarations of unconstitutionality in the chosen jurisdictions. This raises the question of whether to give priority to theoretical considerations, or data and trends discerned from the jurisdictional study. This section accounts briefly for the thesis' approach to this methodological question.

Endicott has addressed this problem through the simple question 'is it a concern for the philosopher that they should surprise lawyers?'[31] One could equally ask this question in the inverse: 'is it any concern for the comparative lawyer that they should surprise philosophers?' Endicott, broadly following Hart, suggests that the aim of a legal theorist should be to account for legal practice as understood by people engaged with it. In other words, the data produced from the jurisdictional survey take priority. It is not that the data are wrong in light of the theory; rather, the theory must generally account for phenomena described by the data. However, this is not to say that any theoretical content articulated in the thesis that surprises a lawyer is to be discarded.

In this thesis I use theoretical analysis to posit possible models rationalising the practice of judicial review of legislation in the chosen jurisdictions. This is the bit that should not surprise lawyers. Descriptively, my use of legal theory aims to ra-

[31] Timothy Endicott, 'Raz on Gaps—The Surprising Part' in Lukas Meyer, Stanley Paulson and Thomas Pogge (eds), *Rights, Culture and the Law: Themes from the Legal and Political Philosophy of Joseph Raz* (OUP 2003) 102–04.

tionalise current practice. However, theoretical criticism of that framework, or the models generated by that framework, is a different consideration. I need not endorse every model produced under the framework, and if criticism I make of them is apt to provoke lawyerly surprise then this is not a concern. In short: the degree to which my theoretically-derived models accurately map to legal practice is an aspect of the work that treats the jurisdictional data as paramount. The degree to which these theoretical models are then sound or flawed according to different metrics is an aspect of the work to which the internal perspective of the legal practitioner is less relevant.

1.3. STRUCTURE OF THE THESIS

1.3.1. Part One: Introduction

This part contains a general introductory chapter, which outlines the subject, approach, and methodology of the thesis, and chapter 2 'Textual Bases for Judicial Review'. This second chapter catalogues and compares the constitutional bases for judicial review in the surveyed jurisdictions. This provides an early understanding of the relative similarity, or difference, between the constitutional baselines of the studied jurisdictions on the question of judicial review of legislation. It is also a useful frame of reference for determining the extent to which textual constraints (where applicable) affect practice. This theme is revisited towards the end of the thesis.

1.3.2. Part Two: Declarations of Unconstitutionality in Practice

This part has five chapters, each of which provides an outline of prominent features of the practice of unconstitutionality in the jurisdictions chosen above. These chapters thus set out the groundwork for a comparative legal analysis of the laws of Ireland, Canada, the United States, India and South Africa. This analysis generates points of data and analysis that are used in Part III to synthesise, and critically reflect on, understandings of unconstitutionality.

Some significant findings of this Part include: (1) the commonality of unconsti-

tutionality entailing invalidity and voidness *ab initio*, or being considered 'objective'; (2) the extent to which the severity of unconstitutionality has been mollified in each jurisdiction; (3) the separation in some jurisdictions, for example the United States, between the court declaring a piece of legislation facially invalid and the court declaring the legislation inapplicable to a specific set of facts; and (4) where there is express textual provision for the retention of pre-constitution law, the difference in treatment between pre-constitution law that is found unconstitutional, and post-constitution law that is found unconstitutional.

1.3.3. Part Three: Theoretical Analysis

This part contains three chapters. The first of these chapters, 'A Theoretical Framework for Unconstitutionality', presents models of unconstitutionality centred around three questions: the derivation question, the effects question, and the temporal question. In order, these questions are

> *Derivation:* Whence does unconstitutionality derive? In virtue of what exercise of power, or failure of the exercise of power, is unconstitutionality established?
>
> *Effects:* What does unconstitutionality do? In what way does unconstitutionality modify legal norms?
>
> *Temporal:* When does unconstitutionality start, and for what duration(s) is it effective?

Answers to these questions are drawn from theoretical literature and the jurisdictional surveys in Part II. The derivation question can be answered by characterising unconstitutionality as requiring an intervention in the form of a judicial order, or as arising from a failed exercise of the legislative power. These provide different constitutive accounts of how unconstitutionality comes to be. The effects question can be answered by characterising unconstitutionality as rendering a legal norm either invalid (not a member of the legal system) or inapplicable (inapt to

justify legal results in the cases it covers). The distinction between these properties is described more fully in the chapter. Finally, the temporal question can either be answered by holding that unconstitutionality begins from the point of its apprehension, or from some point in the past, and that it covers a duration that is both retrospective and prospective, or one that is prospective only.

From these questions, general models of unconstitutionality are constructed by demonstrating how the answers may combine. These models are then mapped back on to the jurisdictional practices that they describe. Accounting for these practices in this theoretical framework makes certain trends among the jurisdictions clearer.

The next chapter in this part, 'The Place of Unconstitutional Law in the Legal System', supplements the previous chapter by substantiating an account of how unconstitutional law can still meaningfully relate to legal systems. This is intended to explore alternatives to the occasionally articulated view that unconstitutional law does not 'exist'. The chapter presents a set-theoretic model of legal systems that shows how unconstitutional law can still be a member of certain subsets of a legal system, even if it is not a member of the valid set of norms within that system. This understanding of unconstitutional law challenges the primacy of validity as an immutable and controlling fact of unconstitutionality, the inexorable consequences of which must sometimes be mitigated or circumvented to prevent catastrophic consequences.

The third chapter of this part, 'Flaws in Existing Theories of Unconstitutionality', subjects some of the major models of unconstitutionality to critical scrutiny. In particular, it criticises the frequently-held view that unconstitutional legal norms are 'void *ab initio*'. First, this view is criticised for requiring a view of law as being objective, insofar as it requires that correct legal answers pre-exist cases in which gaps in the law arise. It is argued that this is implausible both in general, and for constitutional law in particular (that is, even if *some* law is objective, constitutional law is unlikely to fall in this category). The chapter also suggests that the void *ab initio* view is incompatible with a view of legal answers being partly subjectively determined through adjudication. Subjectivity here refers to the idea that legal an-

swers do not pre-exist cases in which legal gaps arise, and so those gaps must be filled through some creative process. Since creativity is particular to individuals, it will be subjective in some measure. It therefore seems that the void *ab initio* view faces difficulties on both the understanding that law is objective and the understanding that it is subjective. This makes this view difficult to maintain in a theoretically satisfying way.

Suspended declarations of unconstitutionality and other remedies that mitigate the harshness of voidness *ab initio* are also subjected to critical scrutiny. These criticisms include the suggestion that such remedies can remove incentives to engage in constitutional litigation, and that they present compatibility issues with the rule of law where an apex court, without much further qualification, effectively condones a law that it knows (and has exposed to be) constitutionally flawed.

1.3.4. Part Four: Conclusion

The fourth and final part, 'Conclusion', uses the analysis from Part III to make observations and recommendations on the practice of unconstitutionality. This part contains just one chapter, 'Implications for Constitutional Theory and Practice'. This chapter argues for a particular set of answers to the questions identified in the 'A Theoretical Framework for Unconstitutionality' chapter. This is the set that has the most surmountable difficulties, taking into account the analysis in the 'Flaws in Existing Theories of Unconstitutionality' chapter, and it is the one that provides the most flexible baseline position.

The chapter makes some further recommendations on how the theory and analysis in this thesis can be used to supplement novel literature and approaches in the field of public law remedies. In particular, the distinction between validity and applicability can be used to view the suspended declaration of unconstitutionality in a new light, and to address some of the criticisms that are levelled at it in Part III.

The chapter also makes observations on the implications of the finding of this thesis for comparative constitutional law more broadly. It does this by disputing the idea that unconstitutionality is uniform or monosemous. This thesis demonstrates

both empirically (in Part II) and theoretically (in Part III) how unconstitutionality differs from place to place. The use of foreign law by judges and academics in this area therefore poses certain risks and challenges. These risks are seldom perceived due to the under-theorisation of unconstitutionality.

2 Constitutional Bases for Judicial Review

CHAPTER OVERVIEW

2.1. Introduction 26

2.2. Textual Foundations 27

2.2.1. Concrete Review 27

2.2.1.1. Fundamental Rights Review 27

2.2.1.2. Federal Review 34

2.2.1.3. Observations 37

2.2.2. Abstract Review 39

2.2.2.1. Ireland 39

2.2.2.2. Canada 41

2.2.2.3. South Africa 41

2.2.2.4. India 42

2.2.3. Constitutional Conditions for a Bill to Count as Law 42

2.2.3.1. United States 43

2.2.3.2. Ireland 43

2.2.3.3. Canada 44

2.2.3.4. India 45

2.2.3.5. South Africa 46

2.3. Non-Textual Foundations 47

2.4. Conclusions 50

2.1. INTRODUCTION

The goal of this chapter is to ascertain the constitutional foundations of the practice of judicial review of legislation in the jurisdictions selected for this study. As each of the jurisdictions has a written Constitution, in most cases this requires little more than pointing to the relevant constitutional article or articles. The initial discussion outlines the textual basis, if there is one, for judicial review in a jurisdiction. This is then divided into provisions on concrete and abstract review. As will become clear, the general textual starting point of each jurisdiction is to find unconstitutional instruments invalid and void.

Constitutional provisions on when a Bill is deemed to become law are also noted. These are relevant to this thesis because unconstitutionality is intimately related to the circumstances in which a law is supposedly not actually law. That is, unconstitutionality could be considered a total failure of legislation. Alternatively, unconstitutionality could relate to the circumstances in which a law fails to acquire juridical or normative significance. Such law would be law, but it would be law that no-one (including a court) need adopt as a reason for action. It is common that signature by a President or some other designated official is required to make a Bill become law. Constitutional assent provisions such as these provide a foundation in these systems for a separation between a law's validity and its applicability.[1] Unconstitutionality must target one of these things; to maintain that a law is unconstitutional could be to maintain either that it was not validly passed, or that it is not applicable in a case or class of cases.

As will later become apparent, there are some significant differences between how each jurisdiction ultimately handles the practice of unconstitutionality, but it is important to note how they all have a relatively common starting point from which differences later emerge. There are, in general, high degrees of similarity among the constitutional provisions discussed below, except for the United States, which has

[1] This distinction will become important in chapter 8.

no textual provision for judicial review of legislation.

2.2. TEXTUAL FOUNDATIONS

Each of the jurisdictions in this study bar one—the United States—has a textual basis grounding constitutional review of legislative compliance with fundamental rights. In federal systems there is usually a separate basis for review of state legislative compliance with the federal constitution.

All the jurisdictions in this study have a textually grounded post-promulgation assent procedure for legislation. This section gives a brief overview of the textual provisions that pertain to abstract and concrete review, where applicable, and the assent procedures for a Bill to be deemed law.

2.2.1. Concrete Review

Concrete review provisions can be divided roughly into two camps. First, there are provisions that allow a court to review legislative enactments for their compatibility with constitutional rights (in a federal system this review will, naturally, be targeted at federal legislation). Second, there are provisions that allow federal courts to review state legislative enactments for breach of devolved competence under the federal constitution. Where applicable to the jurisdictions in this study, I outline these below. Fundamental rights review will be the chief focus of this thesis, as it is common to all jurisdictions, whereas federal review is absent in one (Ireland).

2.2.1.1. Fundamental Rights Review

There are some interesting discrepancies in the wording used in constitutional instruments that provide for constitutional review examining compliance with fundamental rights. As we shall see in chapters 4 to 7, each of these provisions has, at one time or another, been used to support findings of voidness or invalidity. The extent to which this is compelled by the wording of individual Constitutional texts in each case is sometimes questionable. Of course, constitutional law is sometimes more

than simply the constitutional text.[2] For present purposes all that matters is that the constitutional text is a starting point, even if a holistic view of the constitutional law of a jurisdiction ultimately better supports a practice other than that immediately suggested by the text.

The strongest statement regarding the effect of concrete review for breach of fundamental rights is that contained in the Indian Constitution of 1950. Of all the instruments of relevance to this study, only India's Constitution mentions the word 'void' expressly in Article 13:

> (1) All laws in force in the territory of India immediately before the commencement of this Constitution, in so far as they are inconsistent with the provisions of [Part III – Fundamental Rights], shall, to the extent of such inconsistency, be void.
>
> (2) The State shall not make any law which takes away or abridges the rights conferred by [Part III] and any law made in contravention of this clause shall, to the extent of the contravention, be void.

It is interesting to note that the first clause of this Article deals with law that pre-dates the Constitution,[3] and the second deals with post-constitution law. Each type of law may contradict the Constitution differently. Pre-constitution law may be 'inconsistent' with the Constitution, but post-constitution law may 'contravene' it. This very tacitly suggests a difference between pre- and post-constitution law that is borne out in both India and Ireland.[4]

Article 32 provides for the jurisdiction of the courts of India as regards judicial review of legislation:

> (1) The right to move the Supreme Court by appropriate proceedings for the enforcement of the rights conferred by [Part III – Fundamental Rights] is guaranteed.
>
> (2) The Supreme Court shall have power to issue directions or orders or writs, including writs in the nature of *habeas corpus*, *mandamus*, prohibition, *quo warranto* and *certiorari*, whichever may be appropriate, for the enforcement of any of the rights conferred by this Part.

---

[2] For a discussion of this point, see Oran Doyle, 'Conventional Constitutional Law' (2015) 38 DULJ 311.

[3] In the Indian context, such law would encompass law enacted by the Imperial Legislative Council and, after the Indian Independence Act 1947, the Constituent Assembly of India.

[4] See further chapters 5 and 6.

This article makes it clear that the Supreme Court is the court with overall supervisory jurisdiction in guaranteeing the protection of fundamental rights under the Indian Constitution. Other, more general, jurisdiction provisions supplement this guarantee, such as those provided for in Articles: 131,[5] 132(1),[6] 133(1)[7] and 134.[8]

Constitutional texts in other jurisdictions use formulations that do not themselves employ the word 'void'; however, in most cases these alternative formulations have been taken to achieve more or less identical effect in practice. Take, for instance, judicial review of legislation under the Charter of Rights and Freedoms 1982 in Canada. This review has its base in section 52(1) of the Canadian Constitution Act 1982, which states that:

> The Constitution of Canada is the supreme law of Canada, and any law that is inconsistent with the provisions of the Constitution is, to the extent of the inconsistency, of no force or effect.

Although this provision makes no mention of 'voidness', it has been interpreted as effectively imputing this concept regardless.[9] The Canadian formulation is quite flexible in this way, since 'of no force or effect' is not a legally technical term like 'voidness' or 'invalidity'. This gives the judiciary more scope to supply more nuanced or technical readings that would still fall within the broad ambit cast by this provision.

As regards the jurisdiction of the courts in such review, section 24(1) provides that:

> Anyone whose rights or freedoms, as guaranteed by this Charter, have been infringed or denied may apply to a court of competent jurisdiction to obtain such remedy as the court considers appropriate and just in the circumstances.

It is not immediately obvious based on the Constitutional text alone what 'court

---

[5] Providing for the original jurisdiction of the Supreme Court

[6] Appeal from the High Court on certified point of law.

[7] Appeal from the High Court in civil matters.

[8] Appellate jurisdiction in criminal matters.

[9] This is discussed further in chapter 4.

of competent jurisdiction' means. The Supreme Court of Canada provided some clarification on this point in *R v Rahey*, wherein it was made clear that at least provincial superior and appellate courts, and courts created by the federal government, will qualify as courts of competent jurisdiction for the purposes of section 24(1).[10] This thesis focuses primarily on the federal Canadian courts and so further analysis of this point is not required.

The Irish Constitution, in a manner more similar to the Indian than Canadian Constitution, continues the trend of using relatively technical terms when it speaks of 'invalidity'. The relevant text is given by Article 15.4:

> 1° The Oireachtas [Parliament] shall not enact any law which is in any respect repugnant to this Constitution or any provision thereof.
> 2° Every law enacted by the Oireachtas which is in any respect repugnant to this Constitution or to any provision thereof, shall, but to the extent only of such repugnancy, be invalid.

Like its Indian counterpart, the Irish Constitution has a specific provision for the continuance of laws in force that pre-date the Constitution.[11] Laws applicable in the Irish Free State, which preceded the current state of Ireland, are retained by operation of Article 50, which reads as follows:

> 1. Subject to this Constitution and to the extent to which they are not inconsistent therewith, the laws in force in Saorstát Éireann [the Irish Free State] immediately prior to the date of the coming into operation of this Constitution shall continue to be of full force and effect until the same or any of them shall have been repealed or amended by enactment of the Oireachtas.
> 2. Laws enacted before, but expressed to come into force after, the coming into operation of this Constitution, shall, unless otherwise enacted by the Oireachtas, come into force in accordance with the terms thereof.

The effect of Article 50 is to carry over the pre-1937 laws of the Irish Free State into the Republic of Ireland.[12] If a pre-1937 law is found inconsistent with the 1937

---

[10] *R v Rahey* [1987] 1 SCR 588 (SC).

[11] See generally Gerard Hogan and Gerry Whyte, *J M Kelly: The Irish Constitution* (4th edn, Tottel Publishing 2003) 8.2.82 – 8.2.114 for an analysis of this Article that informs my analysis below.

[12] This provision duplicated a provision that had existed in Article 73 of the Irish Free State Constitution of 1922. The purpose of both provisions was the same: to preserve the pre-1922 law and, in the case of the 1937 Constitution, to preserve whatever law had been enacted between 1922 and 1937.

Constitution then it is deemed never to have carried over and thus never had effect in the Republic of Ireland. The language of 'inconsistency' here—rather than 'invalidity' or 'repugnancy'—is employed because the laws in question are assumed to be valid by reference to the power of the Parliament that enacted them.[13] The question, rather, is whether they continue to be *applicable* in the Republic of Ireland.[14] It is interesting to note that the Irish Courts accept a view similar to hard positivism here: the validity of a law is to be determined by reference to its sources, not its merits.[15]

The jurisdiction of the Superior Courts in respect of review under both Article 15.4 and Article 50 is provided by Article 34.3.2°:

> Save as otherwise provided by this Article, the jurisdiction of the High Court shall extend to the question of the validity of any law having regard to the provisions of this Constitution, and no such question shall be raised (whether by pleading, argument or otherwise) in any Court established under this or any other Article of this Constitution other than the High Court, the Court of Appeal or the Supreme Court.

The language of 'invalidity' is also employed in the South African Constitution. The textual ground providing for judicial review of legislation, and its attendant

---

[13] 'I think it is clear from these various provisions of the Constitution that the laws referred to are statutory provisions, as distinct from non-statutory law, and the validity of any statute can only be examined in the light of the powers of the parliament that enacted it'. *The State (Sheerin) v Kennedy* [1966] IR 379 (SC) 386 (Walsh J).
Of course, this position led Walsh J to difficulty when it came to the matter of common law rules being carried over. Initially, it appeared that he was willing to uphold the strict logic of his view in *Sheerin*: 'I do not think that Article 50 … refers to any law other than statute law, and in my view the text of Article 50 makes that clear.' *Gaffney v Gaffney* [1975] IR 133 (SC) 151. This view did not receive much support in later cases: *McKinley v Minister for Defence* [1992] 2 IR 333 (HC) 347. More recently, the Supreme Court revisited the same rule at issue in the *Gaffney* case just cited and held that it had not been imported by Article 50: *W v W* [1992] 2 IR 476 (SC). This holding presumes that Article 50 was applicable to the common law rule in the *Gaffney* case to begin with and, thus, seems like an implicit rejection of Walsh J's strict reading of Article 50.

[14] This statement was expressed in *The State (Sheerin) v Kennedy* (n 13). See also: 'No question could arise as to the power of the Parliament of the United Kingdom or the Oireachtas of Saorstát Éireann to enact such measures: if they have ceased to be of effect in our law, it is because, although validly enacted by the legislature in question, the impugned provision is inconsistent with the provisions of the Constitution and, hence, did not survive the enactment of the Constitution by virtue of Article 50.1.' *The People (DPP) v MS* [2003] 1 IR 606 (SC) 614.

[15] 'The filtering process provided by Article 73 of the 1922 Constitution (like the comparable provision in the 1937 Constitution) related to the content of the law and not its source.' *Geoghegan v Institute of Chartered Accountants in Ireland* [1995] 3 IR 86 (HC) 95 (Murphy J).

remedies, is section 172(1), the text of which reads as follows:

> 1. When deciding a constitutional matter within its power, a court
>   a. must declare that any law or conduct that is inconsistent with the Constitution is invalid to the extent of its inconsistency; and
>   b. may make any order that is just and equitable, including
>   i. an order limiting the retrospective effect of the declaration of invalidity; and
>   ii. an order suspending the declaration of invalidity for any period and on any conditions, to allow the competent authority to correct the defect.

The allocation of jurisdiction to review legislation for constitutional compliance in South Africa is a little more complex than in the other jurisdictions considered here, owing to South Africa's system of centralised constitutional review. Regarding concrete judicial review, it is worth noting that the South African Constitutional Court has primacy in constitutional matters, but other courts play a significant role in the development of constitutional jurisprudence as well. The Constitutional Court has exclusive jurisdiction over certain matters. Rautenbach and Malherbe have summarised these as follows:[16]

> (a) Disputes between national or provincial organs of state concerning the status, powers or functions of those organs.
> (b) Constitutionality of parliamentary or provincial bills (on reference from the President or premier).
> (c) Constitutionality of parliamentary or provincial act at the request of the President or premier.
> (d) Constitutionality of amendments to the Constitution.
> (e) Failure of the Parliament or President to discharge a constitutional obligation, per section 167(4) of the Constitution.
> (f) Certification that a provincial constitutional text, or amendment thereof, complies with the provisions of the Constitution.

Because the Supreme Court of Appeal has jurisdiction to decide appeals on any matter[17] it has competence to hear all constitutional matters that do not fall within

[16] IM Rautenbach and EFJ Malherbe, *Constitutional Law* (LexisNexis Butterworths 2004) 230–31.

[17] Constitution of South Africa 1996, s 168(3).

the exclusive jurisdiction of the Constitutional Court, as outlined in the list immediately above. Similarly, the High Court has jurisdiction to hear all constitutional matters that fall outside the exclusive domain of the Constitutional Court.[18]

It is worth noting that section 172(1) seems to assume that the default remedial response to unconstitutionality is a fully retrospective declaration of invalidity. This is the most natural reading of section (1)(a). It is also the most obvious way to animate the qualifying gloss given to that provision in subsections (1)(b)(i) and (1)(b)(ii). These subsections presume a requirement to mollify harsh applications of a default application of retrospectivity; the orders referred to in those subsections would be of limited additional value if declarations of invalidity were not of immediate retrospective effect. The mandatory nature of subsection (1)(a) has not escaped the notice of the Court, either, and it remarked in *Dawood* that 'a court is obliged, once it has concluded that a provision of a statute is unconstitutional, to declare that provision to be invalid to the extent of its inconsistency with the Constitution'.[19] This general understanding of section 172(1) is also reflected, as we shall see in chapter 7, in the South African courts' 'doctrine of objective invalidity'.

Another important point to note about Article 172(1) is that it specifically uses the word 'invalid'. This is particularly salient when considered in contrast to section 149, which deals with conflicts between federal and provincial legislation:

> A decision by a court that legislation prevails over other legislation does not invalidate that other legislation, but that other legislation becomes inoperative for as long as the conflict remains.

That invalidation is deemed inappropriate for inter-legislation conflicts, but appropriate for conflicts between the constitution and legislation, denotes some clear special import and meaning for the concept of validity in the South African constitutional order. This loosely tracks the same distinction between the treatment of pre- and post-constitution law in the Irish and Indian Constitutions, discussed above.

---

[18] Constitution of South Africa 1996, s 169(a).

[19] *Dawood v Minister of Home Affairs* [2000] ZACC 8, (2000) 3 SA 936 [59].

Finally, the South African Constitution also makes provision for pre-constitution law in Schedule 6, Item 2, which reads as follows:

> (1) All law that was in force when the new Constitution took effect, continues in force, subject to
>
> a. any amendment or repeal; and
>
> b. consistency with the new Constitution.
>
> (2) Old order legislation that continues in force in terms of subitem (1)
>
> a. does not have a wider application, territorially or otherwise, than it had before the previous Constitution took effect unless subsequently amended to have a wider application; and
>
> b. continues to be administered by the authorities that administered it when the new Constitution took effect, subject to the new Constitution.

As shall be seen in chapter 7, through an examination of the doctrine of objective invalidity, the South African Constitution's concept of validity is a part of a general view that legal validity is a property that is regulated on the law's own terms and not at the discretion of officials within the system. Where section 149 leaves open the possibility of the 'revival' of legislation that conflicts with higher-order legislation, in a process similar to the revival of unconstitutional laws in the United States,[20] the contradistinction of the language in sections 149 and 172(1) seems to foreclose on this possibility with respect to legislation that impugns the Constitution.

2.2.1.2. Federal Review

The federal division of the United States is accounted for in Article I of the Constitution. Section 8 of that Article specifically identifies several areas that are the sole province of the Congress:

> 1: The Congress shall have Power To lay and collect Taxes, Duties, Imposts and Excises, to pay the Debts and provide for the common Defence and general Welfare of the United States; but all Duties, Imposts and Excises shall be uniform throughout the United States;
>
> 2: To borrow Money on the credit of the United States;
>
> 3: To regulate Commerce with foreign Nations, and among the several States, and with the Indian Tribes;

[20] Discussed in chapter 3.

> 4: To establish a uniform Rule of Naturalization, and uniform Laws on the subject of Bankruptcies throughout the United States;
>
> 5: To coin Money, regulate the Value thereof, and of foreign Coin, and fix the Standard of Weights and Measures;
>
> 6: To provide for the Punishment of counterfeiting the Securities and current Coin of the United States;
>
> 7: To establish Post Offices and post Roads;
>
> 8: To promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries;
>
> 9: To constitute Tribunals inferior to the supreme Court;
>
> 10: To define and punish Piracies and Felonies committed on the high Seas, and Offences against the Law of Nations;
>
> 11: To declare War, grant Letters of Marque and Reprisal, and make Rules concerning Captures on Land and Water;
>
> 12: To raise and support Armies, but no Appropriation of Money to that Use shall be for a longer Term than two Years;
>
> 13: To provide and maintain a Navy;
>
> 14: To make Rules for the Government and Regulation of the land and naval Forces;
>
> 15: To provide for calling forth the Militia to execute the Laws of the Union, suppress Insurrections and repel Invasions;
>
> 16: To provide for organizing, arming, and disciplining, the Militia, and for governing such Part of them as may be employed in the Service of the United States, reserving to the States respectively, the Appointment of the Officers, and the Authority of training the Militia according to the discipline prescribed by Congress;
>
> 17: To exercise exclusive Legislation in all Cases whatsoever, over such District (not exceeding ten Miles square) as may, by Cession of particular States, and the Acceptance of Congress, become the Seat of the Government of the United States, and to exercise like Authority over all Places purchased by the Consent of the Legislature of the State in which the Same shall be, for the Erection of Forts, Magazines, Arsenals, dock-Yards, and other needful Buildings;—And
>
> 18: To make all Laws which shall be necessary and proper for carrying into Execution the foregoing Powers, and all other Powers vested by this Constitution in the Government of the United States, or in any Department or Officer thereof.

Although the United States does not provide a textual basis for judicial review itself, these provisions do provide a foothold for an analysis that checks state legislative instruments for compliance with federal constitutional requirements. The non-textual foundations of judicial review in the United States are discussed further below.

The scope of review in Canada was originally purely federal, like that which still operates in Australia.[21] Federal powers are set out in sections 91 and 92 of the

[21] Australia has no Bill of Rights in its Constitution, and so the primary constitutional basis for

Constitution Act 1867, which lay out the exclusive competencies of the federal and provincial legislatures, respectively.[22] As Smith has observed '[t]he inner logic of federalism with its distribution of legislative powers pointed to the need for something like the kind of judicial enforcement that pre-Confederation practice had established'.[23] The Supreme Court of Canada, shortly after its initial establishment under the Constitution Act 1867, showed an immediate willingness to exercise judicial review on federalist grounds.[24]

It is also worth noting that section 2 of the Colonial Laws Validity Act 1865[25] applied to Canada in a way that operated similarly to federal review. This Act provided that:

> Any Colonial Law which is or shall be in any respect repugnant to the Provisions of any Act of Parliament extending to the Colony to which such Law may relate, or repugnant to any Order or Regulation made under Authority of such Act of Parliament, or having in the Colony the Force and Effect of such Act, shall be read subject to such Act, Order, or Regulation, and shall, to the Extent of such Repugnancy, but not otherwise, be and remain absolutely void and inoperative.

In many respects, this subordinated Canadian law to English law similarly to how provincial law is subordinated to federal law, with the distinction that it would be the Privy Council that would exercise this ultimate review function. A significant point of difference is, of course, that no special competence was reserved for Canada, as may happen with provincial legislatures in a federal system. Nevertheless, what is important to note here is the stipulation that conflicts between legislation (rather than legislation and the constitution) could render a legislative instrument *void*.

In India, the provisions dealing with the general competences of the legislature provide for federal judicial review. These are contained in Articles 245, 246 and

judicial review is the maintenance of the distinction between federal and municipal powers. See: Kathleen Foley, 'Australian Judicial Review' (2007) 6 Wash U Global Stud L Rev 281.

[22] Sections 94A and 95 address some limited matters of shared jurisdiction.

[23] Jennifer Smith, 'The Origins of Judicial Review in Canada' (1983) 16 Canadian Journal of Political Science 115, 118.

[24] *Severn v The Queen* [1878] 2 SCR 70 (SCC); *Valin v Langlois* [1879] 3 SCR 1 (SCC).

[25] This Act was disapplied by the Statute of Westminster 1931.

Schedule VII, lists 1, 2 and 3. Article 246 outlines the subject-matter competence of the federal and state parliaments,[26] and Article 245 provides in more general detail for the Parliament and State legislatures as law-making organs.[27]

Finally, the South African Constitution contains detailed provisions on its federal governance system in chapters 3, 4 and 6 of its Constitution. Chapter 3 contains only two sections—40 and 41—that deal with general principles of cooperation between the federal, provincial and municipal levels of government. Chapters 4 and 6 outline in detail the federal legislative power, and the general provincial governmental powers, respectively. These provisions must be read in conjunction with section 149, which deals with review of municipal legislation on federalism grounds.

2.2.1.3. Observations

Taking the federalism provisions first, there is little in the way of significant difference between the jurisdictions on this point save that Ireland does not provide for review on federalism grounds. This is, of course, because Ireland has no federal government. There are no major points of discrepancy between the other jurisdictions. The most significant point is that each of the other jurisdictions provides for this less politically contentious genre of judicial review; while review on federalism

---

[26] (1) Notwithstanding anything in clauses (2) and (3), Parliament has exclusive power to make laws with respect to any of the matters enumerated in List I in the Seventh Schedule (in this Constitution referred to as the "Union List").
(2) Notwithstanding anything in clause (3), Parliament, and, subject to clause (1), the Legislature of any State also, have power to make laws with respect to any of the matters enumerated in List III in the Seventh Schedule (in this Constitution referred to as the "Concurrent List").
(3) Subject to clauses (1) and (2), the Legislature of any State has exclusive power to make laws for such State or any part thereof with respect to any of the matters enumerated in List II in the Seventh Schedule (in this Constitution referred to as the "State List").
(4) Parliament has power to make laws with respect to any matter for any part of the territory of India not included [in a State] notwithstanding that such matter is a matter enumerated in the State List.

[27] (1) Subject to the provisions of this Constitution, Parliament may make laws for the whole or any part of the territory of India, and the Legislature of a State may make laws for the whole or any part of the State.
(2) No law made by Parliament shall be deemed to be invalid on the ground that it would have extraterritorial operation.

grounds questions whether a particular competence has been infringed, a largely technical question, rights-based review may require the judiciary to find legislation unconstitutional on more political or moral grounds. Thus, although the possibility of federal review is important to acknowledge, it is not as stark an illustration of the difficulties that can arise from judicial review and unconstitutionality as rights review.

Turning to the bases of judicial review in the constitutions, there is greater divergence to be observed. Not only are various formulations used—'invalid', 'void', and 'of no force or effect'—but they are also interpreted idiosyncratically in each jurisdiction. Take the phrase 'of no force or effect', for example. This appears in both section 52(1) of the Canadian Constitution Act 1982, and article 50 of the Irish Constitution. However, as shall be seen in the doctrinal analysis in chapters 4 and 5, this has been interpreted to effectively mean 'void' for the purposes of the Canadian provision, and something more like 'inapplicable' for the purposes of the Irish provision.

Even where the constitutional text is clear, as in India's Article 13, it seems not to be necessarily fully determinative of practice. That article makes it clear that the default stance for *both* pre-constitution laws (covered by sub-article (1)) and post-constitution laws (covered by sub-article (2)) should be voidness. However, the practice with respect to pre-constitution laws—specifically the 'doctrine of eclipse'—does not carry through fully on the consequences of this view. This is discussed further in chapter 6.

It thus seems that although constitutional texts can establish a practice of constitutional review, the precise wording of these provisions does little to constrain or delimit practice. There is good evidence to suggest that this is due to a lack of attention to theoretical nuance in this area. Judges may seek instruction from comparative review of jurisdictions without fully appreciating contextual difference.[28]

[28] Compare, for example, the Irish judiciary's assertion of the relative idiosyncrasy of the Irish Constitution in both *Murphy v Attorney General* [1982] IR 241 (SC) and *A v Governor of Arbour Hill Prison* [2006] IESC 45, [2006] 4 IR 88, and the Indian judiciary's heavy reliance on US authority

2.2.2. Abstract Review

Many of the object jurisdictions in this thesis also provide for abstract review of constitutional decisions; that is, review of the provisions either before they are fully enacted or without reference to an actual dispute. In some cases, this can produce additional effects that do not obtain in the case of concrete review. In Ireland, for example, if the Supreme Court finds that a piece of legislation that was referred to it is constitutional, the legislation cannot thereafter be subject to a further challenge under concrete review.

2.2.2.1. Ireland

In Ireland abstract review may be initiated by the President. If there is doubt as to the constitutionality of a Bill (other than a Money Bill or a Bill containing a proposal to amend the Constitution) then it may be referred directly to the Supreme Court. This procedure is outlined in Articles 26.1.1° and 26.2.1°, which read as follows:

> 1. 1° The President may … refer any Bill to which this Article applies to the Supreme Court for a decision on the question as to whether such Bill or any specified provision or provisions of such Bill is or are repugnant to this Constitution or to any provision thereof.
>
> […]
>
> 2. 1° The Supreme Court … shall consider every question referred to it by the President under this Article for a decision, and, having heard arguments by or on behalf of the Attorney General and by counsel assigned by the Court, shall pronounce its decision on such question in open court…

If the Bill is found to be repugnant to the Constitution, then the President must decline to sign it.[29] A further consequence is that the approval of a Bill under the Article 26 abstract review process inoculates the resultant Act from further constitutional challenge. This is as a result of Article 34.3.3°, which reads:

---

in its formative case law on unconstitutionality, notwithstanding the different basis of judicial review in the Indian legal framework when compared to the American one. An account of the development of Indian jurisprudence on this point is provided in chapter 6.

[29] Article 26.3.1° of the Constitution of Ireland.

> No Court whatever shall have jurisdiction to question the validity of a law, or any provision of a law, the Bill for which shall have been referred to the Supreme Court by the President under Article 26 of this Constitution, or to question the validity of a provision of a law where the corresponding provision in the Bill for such law shall have been referred to the Supreme Court by the President under the said Article 26.

This has several deficiencies. First, the Supreme Court may only render one judgment with no dissents,[30] so some judicial concerns may be whitewashed in the final result.[31] In *Re Article 26 and the Housing (Private Rented Dwellings) Bill 1981* the Supreme Court noted the difficulty with the immunisation of an Act from further constitutional challenge, particularly where the challenge is heard in an evidential vacuum where counsel appointed by the Court must advance hypothetical arguments. The Court was concerned that changing scientific evidence could warrant different conclusions on the facts and suggested that this might justifiably reopen questions as to the constitutional validity of an Act.[32] Nevertheless, the rule has persisted.

The idea of 'reopening' an Act for consideration as to its validity is one that does not sit easily with the timeless quality of voidness *ab initio*; as shall be seen in chapter 5, this timeless view of validity was, for a long time, the default rule in Irish courts regarding the effect of unconstitutionality. However, it might be said that the underlying assumption behind Article 26 is consistent with voidness *ab initio* in the following way: if legal validity is some timeless quality then a law either has it (and thus it passes Article 26) or does not have it. Yet, the Supreme Court's comments in the *Housing Bill Reference* suggest that they are reluctant to commit to this theory fully, and it might be observed that the injunction on reopening an Act for consideration as to its validity is restricted to Article 26 references only.

---

[30] Article 26.2.2° of the Constitution of Ireland.

[31] This was also required in cases of concrete review under the old Article 34.4.5°, which was inserted by the Second Referendum (1941) and removed in the Thirty-Third Referendum (2013).

[32] *In Re Article 26 and the Housing (Private Rented Dwellings) Bill 1981* [1983] IR 181 (SC) 186–87.

2.2.2.2. Canada

Canada also has a limited provision for abstract review, in the form of 'advisory opinions' on constitutional issues. These can be referred to the Supreme Court either by the provinces or the federal government. This is provided for by section 53 of the Supreme Court Act 1985, clause S-26.

Notably, the decision given by the court following such a referral is not legally binding, though in practice they have not been deviated from and are usually effective in encouraging policy responses from the government.[33] This is particularly evidenced by the pains to which the Supreme Court went to mollify its grave finding in the *Manitoba Language Rights* reference, which I shall discuss further in chapter 4. This is the most significant exercise of the Canadian reference power with respect to the subject of this thesis.

2.2.2.3. South Africa

In the South African context, both the national[34] and provincial[35] legislatures have the power to refer an Act to the Constitutional Court provided that such a reference is supported by a third (in the National Assembly) or a fifth (in provincial legislatures) of the members, and that the Act was signed into law no more than thirty days before the reference. Additionally, both the President[36] and the premier of a province[37] have the power to refer a Bill not yet passed into law to the Constitutional Court for a decision on its constitutionality. Thus, the South African courts have the facility to review legislative provisions in the abstract both before and after their enactment.

---

[33] The following federal reference questions are a small sample of the effective recent exercise of this jurisdiction: *Reference Re Secession of Quebec* 1998 CanLII 793, [1998] 2 SCR 217; *Reference Re Securities Act* 2011 SCC 66, [2011] 3 SCR 837; *Reference Re Senate Reform* 2014 SCC 32, [2014] 1 SCR 704.

[34] Constitution of South Africa 1996, s 80.

[35] Constitution of South Africa 1996, s 122.

[36] Constitution of South Africa 1996, ss 79, 84(2)(b) and 84(2)(c).

[37] Constitution of South Africa 1996, ss 121, 127(2)(b) and 127(2)(c).

2.2.2.4. India

Finally, India also has an abstract review jurisdiction under article 143 of its Constitution. This procedure allows the President to refer a question 'of such a nature and of such public importance that it is expedient to obtain the opinion of the Supreme Court upon it'. The Court may regulate its own process regarding how it conducts the hearing. Once it has heard the issue, it issues a report to the President as to its opinion; this leaves the final resolution of the reference with the executive power and suggests that the Indian power, as with the Canadian reference power discussed above, is non-binding.

2.2.3. Constitutional Conditions for a Bill to Count as Law

In some countries, provisions outside of those that purely provide for judicial review will be relevant to unconstitutionality. Concrete judicial review requires the court to examine a law as it applies to a real pattern of facts. This requires the object of analysis to be classified as a 'law' in the first place. In some Constitutions there are particular preconditions laid down, and a Bill must satisfy these if it is to become law.[38]

The assent procedures described below are also often used as a baseline for a law being deemed to be 'in force'. This establishes a relationship, and a distinction, between a law being valid (passed by the legislature) and 'in force'. This baseline may be deviated from in certain jurisdictions, leading to the possibility of law that is valid but not in force because it has not met some other condition required of it (usually the signing of some order by a member of the executive branch). This distinction in some respects mirrors the doctrine of *vacatio legis* in civil law systems and will help to ground the distinction between validity and applicability endorsed in chapter 8.

[38] The remarks of Walsh J in the Supreme Court of Ireland are perceptive in this regard: 'In Article 26 of the Constitution, where the term "repugnant" occurs, what is being dealt with is not a law but something that will only become a law if signed by the President'. *The State (Sheerin) v Kennedy* (n 13) 386.

2.2.3.1. United States

Article 1, section 7, clause 2 of the United States Constitution outlines three conditions that a Bill must satisfy for it to count as law:

> Every Bill which shall have passed the House of Representatives and the Senate, shall, before it become a Law, be presented to the President of the United States; If he approve he shall sign it, but if not he shall return it, with his Objections to that House in which it shall have originated, who shall enter the Objections at large on their Journal, and proceed to reconsider it. If after such Reconsideration two thirds of that House shall agree to pass the Bill, it shall be sent, together with the Objections, to the other House, by which it shall likewise be reconsidered, and if approved by two thirds of that House, it shall become a Law. But in all such Cases the Votes of both Houses shall be determined by yeas and Nays, and the Names of the Persons voting for and against the Bill shall be entered on the Journal of each House respectively. If any Bill shall not be returned by the President within ten Days (Sundays excepted) after it shall have been presented to him, the Same shall be a Law, in like Manner as if he had signed it, unless the Congress by their Adjournment prevent its Return, in which Case it shall not be a Law.

The three conditions may be simplified as follows. At first, the Bill must be approved by both Houses of Congress. Thereafter, it must be presented to the President. If the President exercises the veto power, then the Bill may return to the Houses where, if it passes through both by two-thirds majority, it becomes a law. Alternatively, if the President does not sign the Bill after ten days and does not exercise the veto power, then the Bill becomes a law as though Presidential signature had been granted.

Case law has established that a Bill becomes law on the date of its signature by the President.[39] Unless the Bill stipulates a specific time at which it is to come into force, it is deemed to come into force on the day that it becomes law.[40]

2.2.3.2. Ireland

In Ireland, Article 25 provides that the President must sign Bills passed through the legislature (Oireachtas) before they become law:

---

[39] *Gardner v The Collector* 73 US 499, 504 (1868); *Burgess v Salmon* 97 US 381, 383 (1878).

[40] *Matthews v Zane* 20 US 164, 211 (1822).

> 1 As soon as any Bill, other than a Bill expressed to be a Bill containing a proposal for the amendment of this Constitution, shall have been passed or deemed to have been passed by both Houses of the Oireachtas, the Taoiseach [Prime Minister] shall present it to the President for his signature and for promulgation by him as a law in accordance with the provisions of this Article.
>
> 2 1° Save as otherwise provided by this Constitution, every Bill so presented to the President for his signature and for promulgation by him as a law shall be signed by the President not earlier than the fifth and not later than the seventh day after the date on which the Bill shall have been presented to him.

However, this is not a discretionary power on the President's part. Article 13.3.2° makes clear that the President 'shall promulgate every law made by the Oireachtas'. Given its non-discretionary nature, the assent procedure by the President is more of a formality than a significant constitutional safeguard, except in rare cases where the Article 26 reference procedure (described above) is invoked.

The Interpretation Act 2005 provides that an Act is deemed to come into force on the date of its passing,[41] which is the date on which it is signed by the President.[42] However, this general rule may be supplanted by requiring an Act, or individual parts or sections of an Act, to require a Commencement Order to be signed by the relevant Minister before those provisions will come into force. This entails that it is possible for valid legislation in Ireland (ie, those Bills signed by the President) to lack the force of law.[43]

2.2.3.3. Canada

For a Bill to become law in Canada, it must be given royal assent. This process is outlined in sections 55 – 57 of the Constitution Act 1867:

> 55. Where a Bill passed by the Houses of the Parliament is presented to the Governor General for the Queen's Assent, he shall declare, according to his Discretion, but subject to the Provisions of this Act and to Her Majesty's Instructions, either that

[41] Section 16 of the Interpretation Act 2005.

[42] Section 15 of the Interpretation Act 2005

[43] For a recent analysis of commencement provisions centred on Irish law, but with comparative elements, see: Patricia Sheehy Skeffington, 'Commencement Orders: A Creeping Incursion into the Legislature's Law-Making Domain?' (2018) 59 Ir Jur 93, noting that uncommenced law 'exist[s] in a curious ambulatory zone of having the status of law, albeit dormant law which is not in effect for the time being.'

> he assents thereto in the Queen's Name, or that he withholds the Queen's Assent, or that he reserves the Bill for the Signification of the Queen's Pleasure.
>
> …
>
> 57. A Bill reserved for the Signification of the Queen's Pleasure shall not have any Force unless and until, within Two Years from the Day on which it was presented to the Governor General for the Queen's Assent, the Governor General signifies, by Speech or Message to each of the Houses of the Parliament or by Proclamation, that it has received the Assent of the Queen in Council. An Entry of every such Speech, Message, or Proclamation shall be made in the Journal of each House, and a Duplicate thereof duly attested shall be delivered to the proper Officer to be kept among the Records of Canada.

It is interesting to note that these provisions couch themselves in the terminology of a Bill being granted 'force'. Section 5 of the Interpretation Act 1985 also chimes with this view, as it provides that the commencement date of an Act, unless otherwise specified, is the date of royal assent.[44] As in Ireland, this entails that, where an alternative commencement date is specified, Canadian law can be valid but not yet in force.

2.2.3.4. India

Similar to the Canadian provisions above, in section 111 the Indian Constitution provides for an assent power to be exercised by the President:

> When a Bill has been passed by the Houses of Parliament, it shall be presented to the President, and the President shall declare either that he assents to the Bill, or that he withholds assent therefrom:
>
> Provided that the President may, as soon as possible after the presentation to him of a Bill for assent, return the Bill if it is not a Money Bill to the Houses with a message requesting that they will reconsider the Bill or any specified provisions thereof and, in particular, will consider the desirability of introducing any such amendments as he may recommend in his message, and when a Bill is so returned, the Houses shall reconsider the Bill accordingly, and if the Bill is passed again by the Houses with or without amendment and presented to the President for assent, the President shall not withhold assent therefrom.

This process seems to be mostly formal, along the same lines as the Irish provision discussed above, but with an added soft veto power. The legislature seems free to ignore the President's concerns, but it is unlikely that this would be politically

[44] Section 5(2), Interpretation Act 1985.

straightforward without putting some onus on the legislature to further justify its policy position. The Supreme Court of India has also held that a Bill must be made law before the review jurisdiction under Article , discussed above, may be engaged.[45]

Section 5 of the General Clauses Act 1897 establishes that the date of commencement for Acts of the Indian Parliament is, by default, the date of assent. However, the Act, or provisions of the Act, may come into effect on another day by a subsequent activation mechanism where this is specified by the Act.[46]

2.2.3.5. South Africa

Finally, as with all the other jurisdictions considered above, South Africa provides for an assent procedure in section 79 of the Constitution:

> (1) The President must either assent to and sign a Bill passed in terms of this Chapter or, if the President has reservations about the constitutionality of the Bill, refer it back to the National Assembly for reconsideration.
>
> …
>
> (4) If, after reconsideration, a Bill fully accommodates the President's reservations, the President must assent to and sign the Bill; if not, the President must either —
>
> (a) assent to and sign the Bill; or
>
> (b) refer it to the Constitutional Court for a decision on its constitutionality.
>
> (5) If the Constitutional Court decides that the Bill is constitutional, the President must assent to and sign it.

Section 81 goes on to clarify that an Act shall have effect upon its publication pursuant to assent and signature by the President. This reflects the rule in section 13(1) of the Interpretation Act 1957, which is still in force in South Africa. As with the other jurisdictions, this provision allows for Acts to provide alternative arrangements for commencement.

In summary, assent by a designated official is a necessary (but insufficient)[47]

---

[45] *Chotey Lal v The State of Uttar Pradesh* AIR 1951 All 228.

[46] For example, the Code of Criminal Procedure was signed on 25 January 1974, but it only came into force on the 1 April 1974 (per section 1(3) of the Code).

[47] It is insufficient because the Presidential signature cannot override other reasons for which legislation might be invalid, such as unconstitutionality.

condition for a law to be considered valid. In each jurisdiction, this assent procedure is also connected with a law coming into force. However, it is also possible to separate the commencement of a law from this assent procedure, such that the assent procedure performs its validity-conferring function, but the law is not yet applicable to legal cases. Often subsequent commencement provisions will be used where the law will require some institutional framework or support that is not yet in place. It is crucial to note that the separation of commencement and validity here requires that *valid law need not be applicable*. This theme will recur in chapter 8.

2.3. NON-TEXTUAL FOUNDATIONS

The United States is unique among the jurisdictions studied in this thesis for having no written constitutional provision detailing any kind of judicial review. This is not to say that it is completely silent on every cognate topic;[48] however, the restrictions it places textually are limited. It has therefore largely been left to the judiciary, following *Marbury v Madison*,[49] to develop a theory of judicial review and its attendant remedies. The significant case law developments following from this are described in chapter 3.

However, since the United States courts were forerunners in judicial review of legislation not bound by a constitutional text, their practice is particularly idiosyncratic in its theoretical commitments. Hogan has noted that under the theory the federal courts have developed, the United States courts must prefer to *apply* the higher legal instrument (the Constitution) and dis-apply the inferior one (a legislative enactment).[50] This, it must be stressed, is different to invalidating, repealing,

---

[48] For example, while the US Constitution is relatively silent on temporal issues in civil law cases, it has both a federal (art I, § 9, clause 3) and state (art I, § 10, clause 1) prohibition on *ex post facto* criminal liability. This has expressly been held not to apply to civil retroactivity: *Calder v Bull* 3 US (3 Dall) 386, 399–400 (1987). The Supreme Court has further held that these prohibitions apply only to legislation and not to acts of the judiciary: *Frank v Mangum* 237 US 309 (1915). Owing to these restrictions, the case law around retrospectivity concentrates around its role in judicial activity.

[49] *Marbury v Madison* 5 US 137 (1803).

[50] Gerard Hogan, 'Declarations of Incompatibility, Inapplicability and Invalidity: Rights, Remedies

or otherwise derogating the inferior norm.[51] It is occasionally conceptualised as a kind of injunction against enforcement of the statute. This sharply distinguishes the United States from all the jurisdictions discussed above.

Hogan further observes that the *erga omnes* effect of declarations of unconstitutionality in the United States is a result of the doctrine of precedent and not any inherent property of the declaration or Constitution.[52] Courts must apply findings of unconstitutionality to subsequent cases of sufficient factual similarity.[53] This marks a further point of distinction between the United States and the other jurisdictions above. The constitutional texts of the other jurisdictions require the courts to hold unconstitutional legislation *void*, *of no force or effect*, *invalid*, etc. They thus derive *erga omnes* effect from the fact of the unconstitutionality itself. If legislation is invalid, void, or of no force or effect then this is so for *everyone*; there is no such thing as legislation being invalid only for an individual or group of individuals.[54] As the United States federal courts seem to hold that unconstitutional legislation is *inapplicable*, they must supply an alternative basis for applying that unconstitutionality to cases other than the case at bar. This role is instead filled by judicial precedent.

The Supreme Court of the United States has endorsed this view of the law since as early as 1923, when it determined in *Massachusetts v Mellon* that:

> We have no power *per se* to review and annul acts of Congress on the ground that they are unconstitutional. That question may be considered only when the justification for some direct injury suffered or threatened, presenting a justiciable issue, is made to rest upon such an act. Then the power exercised is that of ascertaining and declaring the law applicable to the controversy. It amounts to little more than the negative power to disregard an unconstitutional enactment, which otherwise would stand in the way of the enforcement of a legal right. The party who invokes the power must be able to

and the Aftermath' in Kieran Bradley, Noel Travers and Anthony Whelan (eds), *Of Courts and Constitutions:* Liber Amicorum *in Honour of Niall Fennelly* (Hart Publishing 2014) [16].

[51] I discuss this point in more detail in chapter 8.

[52] Hogan, 'Declarations of Incompatibility, Inapplicability and Invalidity: Rights, Remedies and the Aftermath' (n 50) [16].

[53] This argument was also noted by Melville Nimmer, 'A Proposal for Judicial Validation of a Previously Unconstitutional Law: The Civil Rights Act of 1875' (1965) 65 Colum L Rev 1394, 1415.

[54] It is more likely that if the group affected by the unconstitutionality was this small, the defect could be cured through reading-in or severance, as appropriate.

> show not only that the statute is invalid, but that he has sustained or is immediately in danger of sustaining some direct injury as the result of its enforcement, and not merely that he suffers in some indefinite way in common with people generally. If a case for preventive relief be presented, the court enjoins, in effect, not the execution of the statute, but the acts of the official, the statute notwithstanding.[55]

This same view has been endorsed in other federal cases at both the Supreme Court[56] and the Court of Appeals.[57] It has also been endorsed by some state Supreme Courts.[58] There is, overall, a reasonable degree of academic[59] and judicial[60] support in the United States for the idea that federal courts merely have a power to suspend or enjoin enforcement of an unconstitutional statute, rather than wholesale invalidate or 'repeal' it.

Owing to the generally complex issue of the retrospective application of judge-made law, the federal law on the retrospectivity of constitutional rules and unconstitutionality is complex. The federal courts employ a fine degree of categorisation of cases. One must first determine if the measure to be applied retroactively is judicial in origin (ie, a judgment or precedent) or legislative. If what is in issue is a retrospective piece of legislation, the case is governed by the principles in *The Schooner Peggy*[61] and the modern leading authority of *Landgraf v USI Film Products*.[62] The retrospectivity of legislative provisions is not of much concern in this thesis, but this category is mentioned for the sake of completeness.

---

[55] *Massachusetts v Mellon* 262 US 447, 488. (1923).

[56] *Murphy v National Collegiate Athletic Association* 584 US — (2018) (Thomas J, concurring); *Perez v Ledesma* 401 US 82, 124 (1971); *Steffel v Thompson* 415 US 452, 469 (1974).

[57] *Eubanks v Wilkinson* 937 F2d 1118, 1127 (6th Cir 1991); *Winsness v Yocom* 433 F3d 727, 728 (10th Cir 2006).

[58] *Pidgeon v Turner* 538 SW3d 73, 88 fn 21 (Tex 2017); *Kopp v Fair Political Practices Commission* 11 Cal 4th 607, 624 (1995) (citing: *Dombrowski v Pfister* 380 US 479, 491–92 (1965)).

[59] Nimmer (n 53); David Shapiro, 'State Courts and Federal Declaratory Judgments' (1979) 74 Nw U L Rev 759; Richard Fallon, 'Making Sense of Overbreadth' (1991) 100 Yale L J 876; Stuart Buck and Mark Rienzi, 'Federal Courts, Overbreadth, and Vagueness: Guiding Principles for Constitutional Challenges to Uninterpreted State Statutes' [2002] Utah L Rev 381, 425.

[60] *Jawish v Morlet* 86 A2d 96 (Colum App Ct 1952); *Kopp v Fair Political Practices Commission* 11 Cal 4th 607 (1995).

[61] *United States v The Schooner Peggy* 5 US (1 Cranch) 103 (1801).

[62] *Landgraf v USI Film Producers* 551 US 224 (1994).

If the measure is judicial, including cases that establish or revise constitutional rules, retrospective application will depend on whether the case is civil or criminal. Civil cases heard on diversity jurisdiction are now governed by *Harper v Virginia Department of Taxation*,[63] criminal cases on direct review are governed by *Griffith v Kentucky*,[64] and *habeas corpus* applications are governed by *Teague v Lane*.[65] Each of these strands of case law is explored in greater detail in chapter 3. Fig 2.1 summarises the content of the preceding two paragraphs.

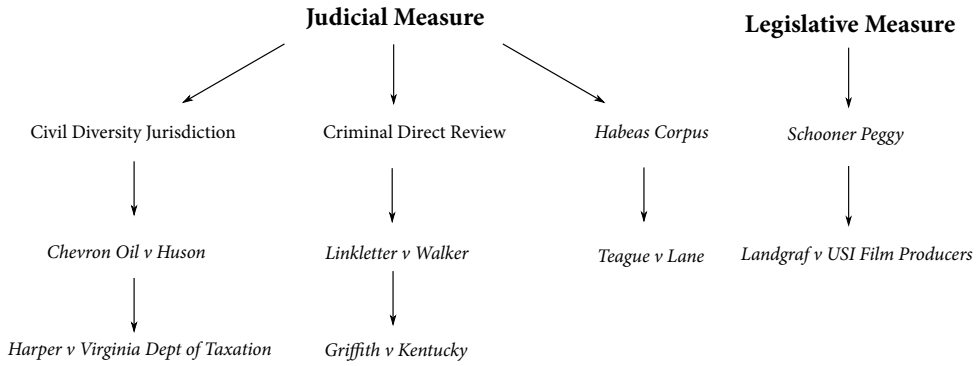


Figure 2.1: Overview of US Law on Retroactivity

2.4. CONCLUSIONS

The principal conclusion to be drawn from this chapter is that constitutional text is rarely determinative of effects of unconstitutional legal norms, even where textual provision is made. Many of the Constitutions that have been quoted in this chapter use phrases such as: 'voidness', 'invalidity' and 'no force or effect'. In reality, such strident views are not borne out, and are usually mollified through remedial doctrines. The chapters to follow will therefore turn to the praxis of unconstitutionality in the courts of the jurisdictions selected for this thesis.

However, textual differences may tacitly point to another conclusion: the instan-

[63] *Harper v Virginia Department of Taxation* 509 US 86 (1993).

[64] *Griffith v Kentucky* 479 US 314 (1987).

[65] *Teague v Lane* 489 US 288 (1989).

tiation of different practices of unconstitutionality. Even if the significance of the text is effaced in the practice of some jurisdictions, it is a significant fact in and of itself that the various constitutional texts studied above attempt to achieve different results following unconstitutionality. Even at this early point in the thesis, therefore, we have reason to suspect that different practices might occur in different jurisdictions. Two further interesting questions arise from this. From a theoretical perspective, it is interesting to contemplate what these practices are and how they might differ. From a comparative perspective, it is interesting to discern whether there is greater or lesser alignment between structurally similar jurisdictions. The individual surveys of each jurisdiction in the chapters to follow will generate further data to be used in interrogating these questions.

A final point to underline is the distinction between a law being valid and a law being 'in force', as discussed in this chapter with reference to constitutional assent procedures. This distinction obtains in every jurisdiction in this thesis, allowing for the possibility that a law can meet all the constitutional requirements to be valid law, but still not be 'in force'. In language that shall be explored and defended later in this thesis, this suggests that the *validity* of a law does not necessarily entail the *applicability* of that law. In other words, just because a law is valid one cannot assume (without further investigation) that it is applicable. If law can be valid but inapplicable, the question of whether the converse holds naturally arises. That is, we might ask whether law can be invalid but applicable. If so, unconstitutional law (invalid law) may not pose so much of a difficulty as is often thought. This theme will be explored in more detail in chapter 8.

Part II.

Declarations of Unconstitutionality in Practice

3 | The United States

CHAPTER OVERVIEW

3.1. Introduction 56

3.2. Doctrines Avoiding or Curtailing Unconstitutionality 57

3.2.1. Avoidance 57

3.2.2. Reading In 58

3.2.3. Severance 59

3.2.4. *Res Judicata* 61

3.3. Early Case Law on Unconstitutionality: Voidness *Ab Initio* 62

3.4. Prospectivity in the Warren Court 67

3.4.1. *Linkletter* and Prospective Invalidity in Criminal Cases . . 67

3.4.2. *Chevron Oil* and Prospective Invalidity in Civil Cases . . . 75

3.5. The Modern Approach to Unconstitutionality and Retroactivity . 80

3.5.1. Retreating from the *Marbury/Norton* View 80

3.5.2. The New Rules on Retrospectivity 83

3.5.2.1. Criminal Cases on Direct Review 83

3.5.2.2. Habeas Corpus Applications 84

3.5.2.3. Federal Civil Disputes 85

3.6. As-Applied and Facial Challenges 87

3.7. Revival of Unconstitutional Statutes 94

3.8. Conclusion 97

3.1. INTRODUCTION

The United States is the first jurisdiction to be surveyed in this part of the thesis. In some respects, the practice of the United States Supreme Court might be thought to have set the leading precedent in the common law world. Courts in both Ireland[1] and India[2] have cited to the United States' jurisprudence on unconstitutionality. The United States also has the distinction of having the longest-running practice of judicial review of any jurisdiction in this thesis, starting with *Marbury v Madison*[3] and continuing to the present day.

It is therefore of little surprise that the jurisprudence of the United States federal courts has shifted so much over time. Periodically, the federal courts have experimented with varying the time periods that are covered by their declarations of unconstitutionality, and even the effect those declarations produce and how that effect might be reversed. This chapter will attempt to trace this long evolution, as well as point out some of the more idiosyncratic features of judicial review in the United States.

The first section, which will be replicated in subsequent chapters containing jurisdiction surveys, will provide a brief overview of doctrines that are used to avoid or minimise a finding of unconstitutionality. These are important contextual features in accounting for the practice of unconstitutionality; for example, if a court has a strong severance doctrine, it may be because it has a harsh unconstitutionality doctrine that it seeks to avoid where possible.

The remainder of the chapter describes the evolution of judicial review of legislation in the United States federal courts. Focus is drawn to the idea of the 'as-applied' challenge to legislation, as well as the revival of unconstitutional statutes. The very possibility of these practices indicates some important implicit theoretical commitments regarding unconstitutionality in the practice of the United States

[1] See chapter 5.

[2] See chapter 6.

[3] *Marbury v Madison* 5 US 137 (1803).

federal courts.

3.2. DOCTRINES AVOIDING OR CURTAILING UNCONSTITUTIONALITY

Each of the jurisdiction surveys in this thesis begins with this section. Doctrines that *avoid* unconstitutionality are those that are designed to allow a court to resolve a case on other grounds. These doctrines allow a court to decline to consider constitutional issues at all. Doctrines that *curtail* unconstitutionality are those that can only be employed once some unconstitutionality has been apprehended by the court, but which help the court to minimise the extent or effect of the unconstitutionality.

3.2.1. Avoidance

The United States has cultivated a doctrine of 'avoidance' for constitutional cases. This is a practice that, as shall be seen in later chapters, occasionally trades under other names, such as 'double construction' in other jurisdictions. The effect of avoidance is that where there are two possible readings of a statute, and one is constitutional and one is not, then the court enforces the constitutional variant.

There are some indications that the requirement that a reading of the legislation be unconstitutional is relatively lax. It may be the case that mere *doubt* as to the constitutionality of a reading may suffice to avoid it where there is an alternative constitutional reading.[4] The intricacies and proper interpretation of the federal courts' avoidance doctrine is not of paramount concern here. What is notable about the practice is the extent to which the federal courts will attempt to avoid unconstitutionality. The permissive application of the avoidance doctrine suggests that unconstitutionality is something that the court will go out of its way to prevent. This, in turn, indicates something about the perception of the effects of unconstitutional-

[4] '[W]here a statute is susceptible of two constructions, by one of which grave and doubtful constitutional questions arise and by the other of which such questions are avoided, our duty is to adopt the latter'. *United States v Delaware & Hudson Co* 213 US 366, 408 (1909). See further: Neal Katyal and Thomas Schmidt, 'Active Avoidance: The Modern Supreme Court and Legal Change' (2015) 128 Harv L Rev 2109.

ity by the courts: unconstitutionality is severe and difficult to constrain effectively and should therefore be employed only as a last resort when alternative methods of resolving the legal difficulty in the case at bar fail.

3.2.2. Reading In

The flexibility seen with avoidance is not generally replicated in other remedies avoiding unconstitutionality in the federal courts. Reading in constitutional language is a good example of this. The Supreme Court has repeatedly stressed its reluctance to interfere with the language chosen by Congress.[5] One exception to this otherwise reluctant attitude, however, is the Court's willingness in equal protection cases to remedy under-inclusive statutes by including the unconstitutionally unprotected groups within the legislative scheme.[6] This exception has been defended on the ground that it better vindicates the will of the legislature, but Fish has correctly observed that this does not seem to adequately single out equality cases.[7]

However, there are other reasons why equality cases might be idiosyncratic; most pertinently, they tend to implicate a limited and clear class of persons, which makes a reading in remedy more clear-cut and restricted in scope.[8] If a law attributes some benefit to men and not women, or Catholics and not Muslims, then it can easily be reshaped simply to include the excluded group in the beneficiary class. This requires less creativity of a judge than other cases might, where there are several permutations of potential remedy that may vary in their extremity.

---

[5] '[I]t is for Congress, not this Court, to rewrite the statute.' *Blount v Rizzi* 400 US 410, 419 (1971). Similar sentiments have been expressed regarding state legislative enactments: 'we will not rewrite a state law to conform it to constitutional requirements.' *Virginia v American Booksellers Association* 484 US 383, 397 (1988).

[6] *Frontiero v Richardson* 411 US 677 (1973); *Weinberger v Wiesenfeld* 420 US 636 (1975); *Califano v Westcott* 443 US 76 (1979); *Blount v Rizzi* 465 US 728 (1984). For general overviews of this area see: Candice Koviacic, 'Remedying Underinclusive Statutes' (1986) 33 Wayne L Rev 39; Eric Fish, 'Choosing Constitutional Remedies' (2016) 63 UCLA L Rev 322, 348–51.

[7] Eric Fish, 'Choosing Constitutional Remedies' (2016) 63 UCLA L Rev 322, 350.

[8] ibid 351.

3.2.3. Severance

The United States also practices severance (sometimes called 'severability').[9] This remedy refers to the wholesale removal of constitutionally offensive words from a statutory text. Following the severance of the unconstitutional language, the statute must satisfy the following three-step test: *(i)* it must function in a manner consistent with the intention of Congress; *(ii)* it must not be legislation that Congress would not have enacted, and *(iii)* it must be capable of independent operation.

This test has evolved steadily since early statements in cases such as *Allen v City of Louisiana*[10] and *Connolly v Union Sewer Pipe Co*.[11] A particularly important foundation was laid down in *Champlin Refining Co v Corp Commission of Oklahoma*. In that case, the Court declared that:

> The unconstitutionality of a part of an act does not necessarily defeat or affect the validity of its remaining provisions. Unless it is evident that the Legislature would not have enacted those provisions which are within its power, independently of that which is not, the invalid part may be dropped if what is left is fully operative as a law.[12]

This statement was recast somewhat in *Alaska Airlines Inc v Brock*, where the Court chose instead to focus on whether the statute after severance could 'function in a *manner* consistent with the intent of Congress'.[13] The Court has revisited this area recently in *Free Enterprise Fund v Public Co Accounting Oversight Board*.[14] In that case, the first two steps of the three-step test described above were endorsed. This reflects the recent trend of the third limb of the test becoming more diluted. In practice, the first two elements of the test—the intent and hypothetical enactment

---

[9] Some detailed recent overviews include: Kenneth Klukowski, 'Severability Doctrine: How Much of a Statute Should Federal Courts Invalidate' (2011) 16 Tex Rev L & Pol 1; Eric Fish, 'Severability as Conditionality' (2015) 64 Emory L J 1293; Fish, 'Choosing Constitutional Remedies' (n 7) 351–53.

[10] *Allen v City of Louisiana* 103 US 80 (1881).

[11] *Connolly v Union Sewer Pipe Co* 184 US 540 (1902).

[12] *Champlin Refining Co v Corp Commission of Oklahoma* 286 US 201, 234–35 (1932).

[13] *Alaka Airlines Inc v Brock* 480 US 678, 685 (1987). Emphasis original.

[14] *Free Enterprise Fund v Public Co Accounting Oversight Board* 561 US 477 (2010).

tests—take precedence.[15]

Most recently, the Supreme Court closely considered its approach to severance in *Murphy v National Collegiate Athletic Association*.[16] This case concerned the validity of the Professional and Amateur Sports Protection Act. The argument was that this act unconstitutionally interfered with the powers of state legislatures by forbidding them from repealing certain laws pertaining to sports betting and barring states from operating their own sports betting operations.[17] The first of these, the proscription of state repeals, was an impermissible infraction on the legislative freedom of the states; however, the second provision merely regulated the type of conduct states could or could not engage in. The majority of the Supreme Court held that this latter provision was not severable from the first, fatally flawed, provision and so both were struck down.

Justice Alito, who delivered the judgment of the Court,[18] thought that it would be improbable to ascribe to the legislature a state of mind where it would have simultaneously thought that it was permissible for states to authorise sports lotteries in private casinos (as it would be once the first provision was struck down), but impermissible for the states themselves to operate such lotteries (as it would be if the second provision stood independently).

Justice Ginsberg delivered the central dissent on severability. She argued that it was not so inconceivable that Congress could have intended the scenario identified by Justice Alito. She also thought that the role of the court when engaging in a severability analysis was to engage in a 'salvage rather than a demolition operation', and so she argued there was more scope to severability in general even outside the bounds of the issue in *Murphy*.

---

[15] Eric Fish, 'Severability as Conditionality' (2015) 64 Emory L J 1293, 1332.

[16] *Murphy v National Collegiate Athletic Association* 584 US — (2018).

[17] Michael Dorf, 'Whither Severability After *Murphy v NCAA*' (*Dorf on Law*, 17th May 2018) ⟨http://www.dorfonlaw.org/2018/05/whither-severability-after-murphy-v-ncaa.html⟩ accessed 11th June 2018.

[18] Justice Thomas joined Justice Alito but wrote a concurring opinion that set out a different view of severability.

It remains to be seen whether the split in the court on severability in *Murphy* will presage further development and changes to the doctrine. Even if the three-step test has been departed from, it seems that the first two steps (both relating to the intention of Congress) will be strictly adhered to and may limit the application of the doctrine.

3.2.4. *Res Judicata*

Res judicata plays an important role in the federal Supreme Court's modern jurisprudence.[19] The policy at work behind *res judicata* is focused on the moral and practical value of legal finality and not on the truth or other merits of the decision. As the US Supreme Court once emphatically stated:

> [T]he maintenance of public order … require[s] that what has been definitely determined by competent tribunals shall be accepted as irrefragable legal truth. So deeply is this principle implanted in [Louisianan] jurisprudence, that commentators upon it have said, the *res judicata* renders white that which is black, and straight that which is crooked.[20]

Notwithstanding this emphatic declaration, there is still some limited scope to circumvent *res judicata* in civil law cases. Rule 60(b) of the Federal Rules of Civil Procedure articulates particular grounds for relief from judgments:

> On motion and just terms, the court may relieve a party or its legal representative from a final judgment, order, or proceeding for the following reasons:
>
> (1) mistake, inadvertence, surprise, or excusable neglect;
> (2) newly discovered evidence that, with reasonable diligence, could not have been discovered in time to move for a new trial under Rule 59(b);
> (3) fraud (whether previously called intrinsic or extrinsic), misrepresentation, or misconduct by an opposing party;
> (4) the judgment is void;
> (5) the judgment has been satisfied, released, or discharged; it is based on an earlier judgment that has been reversed or vacated; or applying it prospectively is no longer equitable; or

[19] See further: Richard Kay, 'Retroactivity and Prospectivity of Judgments in American Law' (2014) 62 American Journal of Comparative Law 37, 51–52.

[20] *Jeter v Hewitt* 63 US 352, 364 (1859). Although the sentiment is directed particularly towards the state of Louisiana, there are no further reasons to suggest that the Court meant some particular comment on that state rather than the doctrine more generally.

(6) any other reason that justifies relief.

Of particular salience here is rule 60(b)(6), which has occasionally been used by the federal courts to relieve parties from final judgments where a change in the law has occurred.[21] There is little case law on this point, however, and so a clear rule is difficult to discern from this limited set of data. All that can be said is that whatever exception rule 60(b)(6) provides, it is to be construed narrowly.

3.3. EARLY CASE LAW ON UNCONSTITUTIONALITY: VOIDNESS *AB INITIO*

The early rhetoric of the US federal courts would suggest that they viewed themselves as invalidating legislation rather than merely disapplying it.[22] Take, for example, the following statement from the seminal judgment of Chief Justice Marshall in *Marbury v Madison*:

> If an act of the legislature, repugnant to the constitution, is void, does it, notwithstanding its invalidity, bind the courts and oblige them to give it effect? Or, in other words, though it be not law, does it institute a rule as operative as if it was law? This would be to overthrow in fact what was established in theory; and would seem … an absurdity too gross to be insisted on.[23]

In this passage, Chief Justice Marshall lends staunch support to the view that an act of the legislature that is legally invalid and void cannot be rendered operative by the courts.[24] This is rendered even clearer by his subsequent argument that in a conflict between legislation and the Constitution, the latter must emerge supreme, which entails rendering the former inoperative:

> So if a law be in opposition to the constitution; if both the law and the constitution apply to a particular case, so that the court must either decide that case conformably

---

[21] *Ackermann v United States* 340 US 193 (1950); *Polites v United States* 364 US 433 (1960).

[22] This view can be contrasted with the recent affirmation by Justice Thomas that this is not the correct view in *Murphy v National Collegiate Athletic Association* (n 16).

[23] *Marbury* (n 3) 177.

[24] However, for a view that *Marbury* did not actually render the provisions of the Judiciary Act 1789 void, see Jonathan Mitchell, 'The Writ-of-Erasure Fallacy' (2018) 104 Va L Rev 933, 964–68.

> to the law, disregarding the constitution; or conformably to the constitution, disregarding the law; the court must determine which of these conflicting rules governs the case. This is of the very essence of judicial duty.
>
> If, then, the courts are to regard the constitution, and the constitution is superior to any ordinary act of the legislature, the constitution, and not such ordinary act, must govern the case to which they both apply.[25]

There is further support for the view that unconstitutional legislation is invalid and void to be found in other classic statements on unconstitutionality. For example, Justice Field stated in *Norton v Shelby County* that: '[a]n unconstitutional act is not a law; it confers no rights; it imposes no duties; it affords no protection; it creates no office; it is, in legal contemplation, as inoperative as though it had never been passed'.[26]

Other strands of case law relatively contemporary to *Marbury* and *Norton* chime with the view that the courts have no role in revising the law. There was strong endorsement of Blackstone's declaratory theory of law-making in the US Supreme Court for some time. This theory holds that the courts do not really 'revise' the law at all; rather, they 'discover' what the law requires, and has always required, and apply the correct rule. The original *locus* of this theory in the jurisprudence of the US Supreme Court is *Swift v Tyson*.[27] In that case, the Court declared that '[i]n the ordinary use of language, it will hardly be contended that the decisions of courts constitute laws. They are, at most, only evidence of what the laws are, and are not, of themselves, laws'.[28] This case was notable for establishing a doctrine of concurrent jurisdiction for state and federal courts,[29] which meant that a federal court could overrule state courts on common law issues (though not issues of statutory construction).[30] Thus, the assertion of the declaratory theory here is perhaps at least partly

---

[25] *Marbury* (n 3) 177–78.

[26] *Norton v Shelby County* 118 US 425, 426 (1886).

[27] *Swift v Tyson* 41 US 1 (1842).

[28] ibid 5.

[29] See: Kermit Roosevelt, 'A Little Theory is a Dangerous Thing: The Myth of Adjudicative Retroactivity' (1999) 31 Conn L Rev 1075, 1084.

[30] *Swift v Tyson* (n 27) 19; *Chicago v Robbins* 67 US 418, 428–29 (1862).

explicable as a deliberately assertive exercise of federal power in its (relatively) early years. Similar statements can also be found in *Gelpcke v City of Dubuque*,[31] where Miller J (dissenting) held that one decision overruling another did not denote 'that the law is changed, but that it was always the same as expounded by the later decision, and that the former decision was not, and never had been, the law'.[32]

These statements endorsing Blackstone were, however, not to last. In *Great Northern Railway Company v Sunburst Oil*[33] Cardozo J held that it was a matter for state courts whether the effects of their judgments would operate prospectively or retrospectively.[34] This was a matter that was up to the 'juristic philosophy' of individual judges; it was not a matter falling within federal jurisdiction. This runs contrary to the *Swift* view that state courts do not create law for that state (as they are merely evidence of law and not themselves law).[35] Although *Sunburst Oil* can be squared with *Swift* by rationalising one case (*Swift*) as about the rules applicable to federal courts, and the other (*Sunburst Oil*) as being about the rules that are applicable to state courts, it is not immediately apparent whether the declaratory theory could permit such a distinction. The declaratory theory is meant to point out a feature of judging in general, and so a distinction between federal and state judges seems irrelevant. Even if it did not entirely dispense with or overrule *Swift*, *Sunburst Oil* represented a clear growing judicial discontent with the declaratory theory. Further discontent with the Blackstonian way of seeing things manifested in several cases after *Sunburst Oil*.[36]

A more decisive death-knell for Blackstone's views in the US federal courts was sounded in *Erie R Co v Tompkins*[37] wherein the Supreme Court declared *Swift* not

---

[31] *Gelpcke v City of Dubuque* 68 US (1 Wall) 175 (1863).

[32] ibid 211.

[33] *Great Northern Railway Co v Sunburst Oil* 287 US 358 (1932).

[34] ibid 364.

[35] See the dissent of Holmes J in *Kuhn v Fairmont Coal Co* 215 US 349, 371 (1910).

[36] See for example: *Mosser v Darrow* 341 US 267 (1951); *James v United States* 366 US 213 (1961); overruling *Commissioner v Wilcox* 327 US 404 (1946).

[37] *Erie R Co v Tompkins* 304 US 64 (1938).

merely undesirable or subject to revision, but flatly unconstitutional. It was recognised that the common law was not *evidenced* by judicial decision but that it *was* judicial decision.[38] This sea change had some repercussions. The declaratory theory obviates retrospective effect by maintaining that the judicial role lies in clarifying (rather than changing) the applicable law at the time the act at issue in the case was committed. In this way, it entails no application of law to acts that pre-date the creation of that law and so avoids retrospectivity.[39] One of the immediate questions following *Erie*, therefore, was how (if at all) the Supreme Court could justify retrospective effect.

The model for retroactive effect first employed by the Supreme Court following *Erie* was what Roosevelt has termed the 'decision-time' model of retrospective effect.[40] This model predicates the retrospective applicability of precedent on the law that was in force *at the time the original decision was rendered*. Thus, a decision would not change the state of the law in the past, and appeal courts would simply apply the law in effect at the time of the original first-instance decision. The federal courts would later prefer (in *Linkletter*) a 'transaction-time' model,[41] which prefers the view that parties should be judged by the law as it stood at the time they undertook the actions that ground the legal dispute. With this as a theoretical starting point, the Court would be more free to employ prospective overruling or other such techniques because they would notionally more fully respect the fairness of treating the parties according to the (perhaps flawed) law of the past when they conducted their affairs. In effect, the *Linkletter* era saw the courts prefer to treat the parties according to the law at the time of their transaction based on legitimate expectations, rather than revise the law applicable to their case, even if the revised law would have been more just.

---

[38] ibid 79. Reed J concurred but felt that the rule was simply erroneous rather than unconstitutional. Butler and McReynolds JJ dissented.

[39] A consequence that Roosevelt notes as 'unpleasantly metaphysical': Roosevelt (n 29) 1083 fn 33.

[40] ibid 1078.

[41] ibid 1079, 1115–17.

The emphatic claims of *Marbury* and *Erie* on the role of the courts in unconstitutionality and revising legal circumstances have been increasingly doubted in more recent years,[42] and, in any case, there are other reasons to believe that the better analysis is that the United States is that unconstitutional legislation is disapplied, rather than invalidated.[43] The most significant aspects of US practice in substantiating this claim are what I term below the 'revival of dead laws' and the practice of differentiating between facial and as-applied challenges. Rhetorically, the US federal courts may employ the language of invalidity or voidness *ab initio*, but their praxis is better squared with the view that federal courts may only suspend or enjoin the enforcement of unconstitutional law.[44] Therefore, even some 'classic' statements on unconstitutionality need to be treated with a degree of caution.[45]

The sequence of this chapter is, however, loosely chronological, at least so far as major developments in the practice of unconstitutionality are concerned. Therefore, before I go on to consider in more detail those aspects of unconstitutionality in the federal courts that would suggest that unconstitutional laws are disapplied, rather than invalidated, other developments must first be accounted for. Most important of these is the development of rules preferring prospectivity in the *Linkletter* case, and the reversal of this standard in a string of subsequent cases.

---

[42] See, for example, *Lemon v Kurtzman* and *Murphy v National Collegiate Athletic Association*, both discussed below.

[43] I am using the term 'disapplied' here as this is the term that I will use in chapter 8. The federal courts do not always use this term; occasionally it is suggested that the courts can enjoin enforcement of a statute, for example. This is effectively disapplying it.

[44] For a discussion that also supports this view, see: Earl Crawford, 'The Legislative Status of an Unconstitutional Statute' (1951) 49 Mich L Rev 645.

[45] However, there are still many references to invalidity. See, for example, Justice Sotomayor's dissent in *Minnesota Voters Alliance v Mansky* 585 US — (2018). As Easterbrook has put it '[t]he Supreme Court has said more times than one can count that unconstitutional statutes are "no law at all".' Frank Easterbrook, 'Presidential Review' (1990) 40 Case W Res L Rev 905, 920.

3.4. PROSPECTIVITY IN THE WARREN COURT

This section accounts for a period of innovation during the Warren Court era where there was a degree of experimentation with a presumption of prospective effect only following unconstitutionality. This practice proved contentious and in a sequence of cases in various areas of substantive law the standard was steadily reversed.

3.4.1. *Linkletter* and Prospective Invalidity in Criminal Cases

Earlier in this chapter, it was observed that the US Supreme Court initially favoured, at least rhetorically, an approach to judicial review of legislation that held unconstitutional laws to be invalid and void. This approach does not allow for what might be called 'true' retroactivity, as it denies there being any actual change to the law of the past. At most, it may only maintain that there is an altered *understanding* on the part of the legal community as to what the law of the past was. The truth or falsity of the legal propositions remains technically unchanged if you grant this theory its assumptions.

For a time, the Court attempted a *volte face* on this point, and entertained a different rule that attempted to address the retrospectivity problem by effectively erasing it; that is, the court maintained that unconstitutional law would be invalidated from the date of its judgment onwards. This animates the problem of retrospectivity, as it will inevitably be the case that justice in some cases will best be served by an extension of the new law into the past circumstances giving rise to the case.

The Warren Court was a significant catalyst in developing the jurisprudence of the Supreme Court on this question.[46] In the landmark case of *Linkletter v Walker*[47] it was held that the court could suspend or withhold the retroactive effect of judicial rulings on unconstitutionality. The question in *Linkletter* was whether a con-

[46] As Roosevelt has put it: 'Before the Warren Court, the question of retroactivity was not found in the case law, for the simple reason that the concept of retroactivity was not there either'. Roosevelt (n 29) 1082.

[47] *Linkletter v Walker* 381 US 618 (1965).

stitutional rule articulated in *Mapp v Ohio*[48] (holding the exclusionary rule under the Fourth Amendment applicable to the states) had retrospective effect. The Court distanced itself from the Blackstonian model of judging, attributing the rival view that judges can make law 'interstitially' to Austin. Of judicial law-making, Austin had claimed that it was a 'childish fiction … that judiciary or common law is not made by [judges], but is a miraculous something made by nobody, existing … from eternity, and merely declared from time to time'.[49] The Court further claimed that the Austinian view was reflected in earlier cases discussed above, such as *Gelpcke v City of Dubuque* and *Sunburst Oil*.[50] Indeed, the leading judge in *Sunburst Oil*, Justice Cardozo, had prior to that case, extra-curially ruminated on the judicial role as follows:

> As the years have gone by, and as I have reflected more and more upon the nature of the judicial process … I have grown to see that the process in its highest reaches is not discovery, but creation; and that the doubts and misgivings, the hopes and fears, are part of the travail of the mind, the pangs of death and the pangs of birth, in which principles that have served their day expire, and new principles are born.[51]

In *Linkletter* the Court confronted more directly this abstract tension between its earlier Blackstonian affiliations as against its relatively more recent 'realist' or 'positivist' leanings.[52] Unimpressed, however, by the inability of either the Austinian or Blackstonian school to provide conclusive guidance to the retroactivity problem, some members of the *Linkletter* court preferred to sidestep this debate entirely.[53] It is perfectly natural that courts should not view themselves as having to arbitrate on vexed theoretical questions to resolve practical disputes; however, it would be equally wrong-headed for courts to simply shun inconclusive theoretical debates in

---

[48] *Mapp v Ohio* 367 US 643 (1961).

[49] John Austin, *Lectures on Jurisprudence* (vol 2, John Murray 1863) 342.

[50] *Linkletter* (n 47) 622–26. Citing: *Gelpcke v City of Dubuque* (n 31); *Sunburst Oil* (n 33).

[51] Benjamin Cardozo, *The Nature of the Judicial Process* (Yale University Press 1921) 166–67.

[52] Ralph Rossum, 'New Rights and Old Wrongs: The Supreme Court and the Problem of Retroactivity' (1974) 23 Emory L J 381.

[53] 'Interesting as the question may be abstractly, this case should not be decided on the basis of arguments about whether judges "make" law or "discover" it when performing their duty of interpreting the Constitution'. *Linkletter* (n 47) 643 (Black J, dissenting).

this way. As Rossum has pointed out: 'While the Court's policy on retroactivity perhaps cannot be determined simply on the basis of jurisprudential considerations, it should at least be informed by them'.[54] This is true particularly in *Linkletter* because in that case it was simply unavoidable that the court would have to commit itself to one position or the other. Though it professed that it would not resolve this debate, implicitly the court committed itself to a position as the standard of general prospective effect adopted in *Linkletter* is consistent only with the 'positivist'/'realist' Austinian school. It would behove courts in this position to at least making a passing attempt at engaging with whatever insights can be gained from different strands of theory rather than incidentally committing itself to one side of the debate as a corollary of a facially agnostic position.

At any rate, two central propositions regarding retrospective effect emerged from the Court's judgment in *Linkletter*:

> (1) a change in law will be given effect while a case is on direct review, and (2) the effect of the subsequent ruling of invalidity on prior final judgments when collaterally attacked is subject to no set 'principle of absolute retroactive invalidity' but depends on a consideration of [vested rights and public policy].[55]

According to condition (1), whether a constitutional rule would have retrospective effect in a given case would turn on whether that case had reached finality or not.[56] This opens itself to a potentially capricious distinction between *habeas corpus* petitioners on the one hand, and defendants pursuing appellate review on the other.[57] The distinction is potentially capricious particularly because the Supreme Court had, not long before *Linkletter*, significantly amplified the significance of constitutional errors in *habeas corpus* applications. In *Fay v Noia* it was held that in certain circumstances *habeas* petitioners could collaterally attack their original

---

[54] Rossum (n 52) 406.

[55] *Linkletter* (n 47) 627.

[56] Collateral attack involves a review of a prior judgment in proceedings other than those that led to that judgment. The most usual collateral attack would be an appeal for a writ of *habeas corpus*. Other types of collateral attack would include an appeal for a writ of *corum nobis*.

[57] Francis Beytagh, 'Ten Years of Non-Retroactivity: A Critique and a Proposal' (1975) 61 Va L Rev 1557, 1601.

convictions on the basis of new constitutional rules not mentioned in the original pleading,[58] and in *Kaufman v United States* it was specifically held that a Fourth Amendment claim of unconstitutional search and seizure, which was not pleaded at trial or on appeal, was nevertheless relevant to a federal *habeas corpus* hearing.[59]

On the facts of the *Linkletter* case, it was decided that non-retroactivity was the sounder course for the rule in *Mapp*. This heralds what Roosevelt terms the 'transaction-time' model for retrospectivity.[60] This model can be contrasted with the 'decision-time' model from the earlier *Erie* era described above. The transaction-time model based retrospectivity on the law applicable at the time the parties to the case undertook the actions generating the legal cause of action. The decision-time model, as the name suggests, applied the law in force at the time judgment at first instance was rendered in the case.

The aftermath of *Linkletter* saw the courts struggle with the boundaries of this new doctrine. In *Stovall v Denno*,[61] which case concerned a *habeas corpus* applicant asserting the retrospectivity of constitutional rules relating to the admission of evidence,[62] the Court reiterated the truism that constitutional rules had to be applied to the cases in which they were announced.[63] There was thus some slight concession to retrospectivity, insofar as the application of a new rule to the petitioner in a case requires retrospective application, given that the facts of the case will, tautologically, have occurred prior to the judgment in the case. The Court also further refined the *Linkletter* approach on the application of new rules to other cases. The new test to determine whether retrospectivity would be appropriate in the context of a collateral attack consisted of three factors for the court to consider, derived from the *Stovall*

[58] *Fay v Noia* 372 US 391 (1963).

[59] *Kaufman v United States* 394 US 217 (1969).

[60] Roosevelt (n 29) 1079.

[61] *Stovall v Denno* 388 US 293 (1967).

[62] Specifically, the rules in *United States v Wade* 388 US 218 (1967); *Gilbert v State of California* 388 US 263 (1967).

[63] *Stovall v Denno* (n 61) 301.

gloss on *Linkletter*. These are summarised by Murphy as follows:[64]

> 1. the new rule's intended purpose and the efficacy of retroactive application of the rule in the furtherance of that purpose (the 'purpose factor');
> 2. the extent of reliance on invalidated precedent by law enforcement authorities (the 'reliance factor');
> 3. the effect of retroactive application on the administration of justice (the 'effect factor').

There were some further teething issues with these new rules however, which led to inconsistent results being achieved in cases, and a sense of unpredictability in what the Court would decide.[65] The Court was notably inconsistent with respect to what the definitive temporal cut-off point would be in denying an applicant relief by retrospective application of a rule. The tripartite test above points to considerations that are unquestionably relevant, but it does little to constrain a court and provide consistency or clarity. O'Sullivan has pointed to at least four different temporal standards that were employed by the Court under this test:[66]

> 1. The rule is denied to those whose convictions have become final prior to the date of the decision that is to be retroactively applied.[67]
> 2. The rule cannot be availed of by those whose trials began before the date of the new decision, whether their convictions are final or not.[68]
> 3. Benefit under the new rule was denied in those cases where the specific constitutional violation occurred before the case laying down the new rule was decided, irrespective of the finality of the conviction.[69]

---

[64] William Mark Murphy, 'The Problem of Unconstitutionality and Retroactivity in Criminal Law: Ireland, the US and Canada Compared' (2007) 42 Ir Jur 63, 85; *Stovall v Denno* (n 61) 297; see also: *Johnson v New Jersey* 384 US 719 (1966); *Tehan v Shott* 382 US 406 (1966).

[65] Rossum (n 52) 402–03; Beytagh (n 57) 1604–05.

[66] Julie O'Sullivan, '*United States v Johnson*: Reformulating the Retroactivity Doctrine' (1983) 69 Cornell L Rev 166, 167–68, fn 8.

[67] *Linkletter* (n 47) (restricting the ambit of the rule in *Mapp* (n 48)); *Tehan v Shott* 382 US 406 (1966) (retrospective application of the rule in *Griffin v California* 308 US 609 (1965)).

[68] *Johnson v New Jersey* 384 US 719 (1966) (rules in *Escobedo v Illinois* 378 US 478 (1964); *Miranda v Arizona* 384 US 436 (1966)); *DeStefano v Woods* 392 US 631 (1968) (rules in *Bloom v Illinois* 391 US 194 (1968); *Duncan v Louisiana* 391 US 145 (1968)).

[69] *Stovall v Denno* (n 61) (rules in *United States v Wade* 388 US 218 (1967); *Gilbert v State of California* 388 US 263 (1967)); *United States v Peltier* 422 US 531 (1975) (rule in *Almeida-Sanchez v United States* 413 US 266 (1973)).

4. The rule is denied to applicants in whose cases the illegal evidence was introduced prior to the decision that made such evidence illegal.[70]

This level of *ad hoc* variation is patently unsatisfactory and points to a very underdeveloped approach the the complexities adn nuances of retrospectivity. Other aspects of *Linkletter* were abandoned reasonably quickly. The distinction it drew between final convictions, and those at various other stages of trial or direct review, was abandoned in *Stovall*.[71] The distinction it drew between direct review and collateral attack was abandoned in *Williams v United States*,[72] in favour of a view that both litigants on direct review whose constitutional violations occurred before the date of the new decision and litigants collaterally attacking judgments that had been rendered before the decision that laid down the new rule would be treated the same.[73] The (in some respects arbitrary) distinction between direct and collateral review was now replaced with a rule that excluded both collateral applicants and applicants on direct review whose constitutional rights had been violated prior to the date of the decision that was to be applied retrospectively.

As mentioned above, one constant was that the courts consistently allowed the applicant in the case at bar to benefit from the new rule.[74] The notional constitutional justification for this constant is article III of the United States Constitution, which requires that the courts resolve actual cases and not render constitutional adjudications 'mere dictum'.[75] The Court maintained that if relief was denied to the applicant in the case at bar then this would be to reduce the overruling language to such a dictum. These arguments were not universally accepted, however,

---

[70] *Desist v United States* 394 US 244 (1969) (rule in *Katz v United States* 389 US 347 (1967)); *Fuller v Alaska* 393 US 80 (1968) (rule in *Lee v Florida* 392 US 378 (1968)).

[71] *Stovall v Denno* (n 61) 300.

[72] *Williams v United States* 401 US 646 (1971).

[73] 'Nor [has the Court] accepted as a dividing line the suggested distinction between cases on direct review and those arising on collateral attack'. ibid 651–52. Confusingly, however, the court went on to endorse the distinction again in *United States v Johnson* 457 US 537 (1982).

[74] Justice Douglas passed some remarks on this feature of the case law in *Desist v United States* (n 70) 255–56.

[75] *Stovall v Denno* (n 61) 301; *Desist v United States* (n 70) 254–55.

and some judges[76] and commentators[77] remained troubled by the disequilibrium in treatment between similarly situated defendants other than the petitioner taking the case. As Torcia and King memorably put the problem, the pivotal inquiry with respect to whether an individual is deserving or undeserving of constitutional protection should not be whether they happened to win a race to the courthouse.[78]

Academic criticism of the decision in *Linkletter* and its progeny abounds.[79] Some academics who otherwise supported the result achieved by the court in *Linkletter* decried the abandonment of the Blackstonian theory, suggesting that it may be a myth we simply have to live by[80] (a sentiment that has been echoed elsewhere).[81] But criticism of the case is not purely academic. Justice Harlan wrote trenchant dissents to *Linkletter* shortly after the case was decided, and these dissents formed the philosophical basis of the Supreme Court's later pivot towards retrospective effect.

Justice Harlan was dissatisfied with several aspects of the *Linkletter* doctrine. In particular, he found it intolerable that an applicant on direct review could only benefit from the retrospective application of a rule if it was argued in *their* proceedings.[82] Similarly situated applicants would thus be denied the benefit of the rule. He drew attention to this inequality of treatment between direct review applicants, suggesting that this was antithetical to the very practice of judicial review:

---

[76] *United States v Peltier* (n 69) 543 (Douglas J, dissenting).

[77] Thomas Currier, 'Time and Change in Judge-Made Law: Prospective Overruling' (1965) 51 Va L Rev 201, 201–04; James Haddad, '"Retroactivity Should be Rethought": A Call for the End of the *Linkletter* Doctrine' (1969) 60 The Journal of Criminal Law, Criminology, and Police Science 417, 438; Beytagh (n 57) 1602–04.

[78] Charles Torcia and Donald King, 'The Mirage of Retroactivity and Changing Constitutional Concepts' (1962) 66 Dick L Rev 269, 284.

[79] See generally: Thomas Currier, 'Time and Change in Judge-Made Law: Prospective Overruling' (1965) 51 Va L Rev 201; James Haddad, '"Retroactivity Should be Rethought": A Call for the End of the *Linkletter* Doctrine' (1969) 60 The Journal of Criminal Law, Criminology, and Police Science 417; Pierce Hasler, 'Retroactivity Rethought: The Hidden Costs' (1972) 24 Me L Rev 1; Beytagh (n 57); O'Sullivan (n 66); Roosevelt (n 29).

[80] Paul Mishkin, 'Foreword: The High Court, the Great Writ, and the Due Process of Time and Law' (1965) 79 Harv L Rev 56.

[81] John Finnis, 'The Fairy Tale's Moral' (1999) 115 LQR 170.

[82] Interestingly, this now seems to be the rule in Ireland after *Damache v Director of Public Prosecutions* [2012] IESC 11, [2012] 2 IR 266. See further the discussion in chapter 5.

> Simply fishing one case from the stream of appellate review, using it as a vehicle for pronouncing new constitutional standards, and then permitting a stream of similar cases subsequently to flow by unaffected by that rule constitutes an indefensible departure from [the] model of judicial review.[83]

Rather, Justice Harlan thought that it was the duty of the court to apply the law 'as it is at the time, not as it once was'.[84] In other words, the law from *now* (a better understanding of the Constitution, we assume) should be applied to facts *then*; this is an endorsement of retrospectivity. One might be forgiven for thinking that this logic would also extend to corollary review applicants as well. However, Justice Harlan confessed that the additional issue of what he termed 'choice of law' for *habeas corpus* applicants was more vexed.[85] He thought that while there was a constitutional impetus for federal courts to adjudicate all issues of law on direct review, the same had never been held with respect to *habeas* applications.[86] The scope of *habeas corpus* applications would have to be balanced against considerations not applicable to the direct review context, such as the general finality of criminal convictions.[87]

Justice Harlan saw *habeas* review as serving two broad purposes: *(i)* 'to inquire into every constitutional defect in any criminal trial, where the petitioner remains "in custody" because of the judgment in that trial',[88] and *(ii)* to provide a quasi-appellate review function that requires courts to 'toe the constitutional mark'.[89] In Justice Harlan's view, neither of these functions required the retroactive application of new constitutional rules on *habeas corpus* review, nor did he think that there was any constitutional imperative to always apply existing law to determine whether an individual was in unconstitutional custody.[90]

---

[83] *Mackey v United States* 401 US 667, 679 (1971).

[84] ibid 681.

[85] ibid 682.

[86] ibid 682.

[87] ibid 682.

[88] ibid 685.

[89] ibid 687.

[90] ibid 686–87.

In the 1980s, the *Linkletter* edifice was gradually dismantled by the Supreme Court in a series of cases and replaced with a new model that was inspired by the dissents of Justice Harlan.[91] In *United States v Johnson*[92] the court characterised the rule in *Linkletter* as establishing that 'all newly declared constitutional rules of criminal procedure would apply retrospectively at least to judgments of conviction not yet final when the … rule was established'.[93] The Court acknowledged that some cases had deviated from this principle, but maintained that there were some opinions of justices that were consistent with this rule.[94] In particular, the Court pointed to the earlier dissents by Justice Harlan and concluded that in Fourth Amendment cases, decisions articulating new rules or reversing old rules would be applicable to all non-final convictions at the time the decision was handed down.[95]

The *Linkletter* doctrine was abandoned more completely in *Griffith v Kentucky*.[96] This case is discussed further below under the section addressing the modern US law on retrospectivity in criminal cases. Judge Harlan's dissent in *Mackey* would also prove influential in the modern rule for *habeas corpus* applications as given in *Teague v Lane*[97] (also discussed below).

3.4.2. *Chevron Oil* and Prospective Invalidity in Civil Cases

Early judicial opinion preferred a general rule of non-retroactivity in federal civil cases. However, most of the cases that touched on this point occurred in the context of so-called 'diversity jurisdiction', which enables federal courts to hear civil matters where the parties are denizens of different US states, or non-US states.[98] As *Erie*

[91] O'Sullivan (n 66).

[92] *United States v Johnson* (n 73).

[93] ibid 543.

[94] ibid 545.

[95] ibid 562.

[96] *Griffith v Kentucky* 479 US 314 (1987).

[97] *Teague v Lane* 489 US 288 (1989).

[98] The constitutional basis of this jurisdiction is art III, § 2 of the United States Constitution, and its modern statutory basis is: 30 USC § 1332.

(discussed above) dispelled the idea that there would be a 'federal common law', the federal courts are required to apply state substantive law when exercising such jurisdiction.[99] A frequently recurring issue was thus the temporal boundaries of changes to *state* law affecting federal cases under this jurisdiction. The general trend in the case law was that any change to the interpretation of state law would have prospective application only. The leading statement of this approach was *Rowan v Runnels*, wherein the majority held that:

> [Federal courts] will always feel … bound to respect the decisions of the State courts, and from the time they are made will regard them as conclusive in all cases upon the construction of their own constitution and laws.
>
> But we ought not to give to them a retroactive effect … [f]or, if such a rule were adopted, … the provision in the constitution of the United States, which secures to the citizens of another State the right to sue in the courts of the United States, might become utterly useless and nugatory.[100]

The dissenting opinion noted that this would give the state Constitution (in this case, the State of Mississippi) 'different meanings at different periods of its existence',[101] a theme that is reflected in a more modern context in some judgments of Justice Scalia.[102] Notwithstanding this difficulty, however, the rule in *Rowan* continued to find favour in subsequent diversity jurisdiction cases.[103]

The United States Supreme Court's first major investigation into prospective effect in civil cases outside the context of diversity jurisdiction occurred in *Chevron Oil Company v Huson*.[104] Huson had suffered an injury at work, recovery for which was governed by the Outer Continental Shelf Land Act. The Act specified no limitation period, but it had been assumed in several federal court cases that it

---

[99] The law around the *Erie* doctrine is quite complex, and it is not necessary to explain it in full here. Subsequent landmark cases include *Byrd v Blue Ridge Rural Electric Cooperative* 356 US 525 (1958); *Hanna v Plumer* 380 US 460 (1965); *Gasperini v Center for Humanities* 518 US 415 (1996).

[100] *Rowan v Runnels* 46 US 134, 139 (1847).

[101] ibid 140 (Daniel J, dissenting).

[102] *Harper v Virginia Department of Taxation* 509 US 86, 105 (1993).

[103] *Ohio Life Insurance and Trust Co v Debolt* 57 US 416 (1853); *Gelpcke v City of Dubuque* (n 31); *Douglass v County of Pike* 101 US 677 (1879).

[104] *Chevron Oil Co v Huson* 404 US 97 (1971).

was governed by *laches* and admiralty law.[105] The principle issue in this case was whether the prior Supreme Court case of *Rodrigue v Aetna Casualty & Surety Company*[106] (setting the statutory limit for recovery under the Act as equal to the limitation period of the nearest state) was applicable to the case at bar notwithstanding that it had been decided after the case at bar had commenced. If *Rodrigue* applied, the limitation period would have been set to one year and Huson's claim would be statute barred. In deciding this point, the Court laid down the following criteria for a civil law decision to apply non-retroactively; the decision would have to:

1. establish a new principle of law, either by overruling or rendering decision on an issue for the first time;
2. be a rule such that, by analysis of its prior history, it would hamper its operation to apply it retroactively;
3. not cause inequity by failing to apply retroactively.[107]

In *Chevron Oil* itself eight justices[108] favoured the view that prospective-only application was preferable in this case.[109] Following *Chevron Oil* there was some strict application of the new rule;[110] however, some courts applied a much earlier rule articulated in *The Schooner Peggy*.[111] There was some criticism of the court for its simultaneous adoption of these two rules.[112] This criticism seems misplaced, however, as it misses the point that *Chevron Oil* and *The Schooner Peggy* are directed to different types of retrospective change, as discussed in chapter 2. The *Schooner*

---

[105] *Pure Oil Co v Snipes* 57 F2d 416 (5th Cir 1961); *Movible Offshore Co v Ousley* 346 F2d 870 (5th Cir 1965); *Loffland Bros Co v Roberts* 386 F2d 540 (5th Cir 1968).

[106] *Rodrigue v Aetna Casualty & Surety Co* 395 US 352 (1969).

[107] *Chevron Oil* (n 104) 106–7.

[108] Douglas J did not consider the retrospectivity point

[109] 'Both a devotion to the underlying purpose of the Land Act's absorption of state law and a weighing of the equities requires nonretroactive application of the state statute of limitations here'. *Chevron Oil* (n 104) 109.

[110] *United States v Johnson* (n 73) 563.

[111] *United States v The Schooner Peggy* 5 US (1 Cranch) 103 (1801); applied in: *Gulf Offshore Co v Mobil Oil Corp* 453 US 473, 486, fn 16 (1981); *Saint Francis College v Al-Khazraji* 481 US 604, 608–09 (1987).

[112] John Corr, 'Retroactivity: A Study in Supreme Court Doctrine as Applied' (1983) 61 N C L Rev 745, 796.

Peggy was a case wherein the United States Supreme Court approved *legislative* retroactive change, via an international treaty, that was relevant to a pending case.[113] In that case, a treaty was signed between the United States and France that contained provisions dispositive of the issue. These provisions themselves provided that they were to be applied retrospectively. It has thus been more correctly observed that *The Schooner Peggy*, while sometimes misunderstood as the start of retroactivity jurisprudence generally, was not a case about retroactivity (at least in the same sense as is currently under consideration) at all.[114] The two strands of cases share the commonality that they are about how courts are to react to legal changes, but the origins of those changes are completely different in each. *The Schooner Peggy* is most narrowly about legislative measures that are facially retrospective; *Chevron Oil* is about the retrospectivity of civil rules that are judicial in origin.

Whether there was truly an inconsistency between *Chevron Oil* and *The Schooner Peggy* would prove irrelevant, as the former was to be slowly abandoned by the federal courts. There came further portents of another sea change in the Court's view when, in *Griffith v Kentucky*,[115] it opted for a firmer stance against pure prospectivity in criminal cases and maintained that there was no discretion to limit a new rule to purely prospective application. The Court took the view that temporal limitations of new legal rules (or, at least, constitutional principles relating to criminal procedure in this case) was a legislative function. Granted, this was a criminal case and, at least immediately, this same standard was not extended to civil cases, but it indicated that the court was reconsidering some fundamental issues of retroactivity.

More substantial change in the civil law context began to emerge in a trio of cases that all dealt with restitution under unconstitutional taxation laws. Unfortunately, each of these cases is beset with a plurality of opinions and little by way of clear *ratio*

[113] An issue they revisited more recently in *Landgraf v USI Film Producers* 551 US 224 (1994).

[114] Roosevelt (n 29) 1082–83; Charles Rhodes, 'Over the Threshold of Constitutional Adjudicative Retroactivity' (*SSRN Draft Paper*, 24th May 2016) ⟨https://ssrn.com/abstract=2783954⟩ accessed 18th December 2017, 8, fn 26.

[115] *Griffith v Kentucky* (n 96).

emerges. Certain trends are, however, discernible. The first of these cases, *American Trucking Associations v Smith*,[116] concerned an Arkansas highway tax that was similar in nature to a tax that had been declared unconstitutional in a prior case decided by the Court (*American Trucking Associations v Scheiner*).[117] The Court began by reaffirming *Chevron Oil* as the applicable standard in civil cases.[118] The controlling question was thus whether the Arkansas Supreme Court had applied the *Chevron Oil* criteria correctly. The Court found that in this case non-retroactivity was favourable, and thus the *Scheiner* decision did not apply to the Arkansas highway tax. This result was only achieved by a slim 5:4 majority, however. Four of the justices supported the application of *Chevron Oil*.[119] Four other justices[120] objected to the idea of different sets of law governing the same controversy, with the only relevant variable being when in time those controversies arose.[121] The ninth judge, Justice Scalia, believed that the Arkansas tax in question was constitutional and thus joined the majority. However, he made clear that he was staunchly opposed to the majority's non-retroactivity doctrine, and that he was therefore sympathetic to the minority on this point.[122]

The Court continued to add further gloss to this analysis. In *James B Beam Distilling v Georgia*,[123] the Court considered the idea of selective or modified prospectivity. The issue in *James Beam* was whether the Court's prior holding in *Bacchus Imports v Dias*[124] (declaring a Hawaiian excise tax on alcohol unconstitutional) could be retroactively applied to a similar Georgian provision. The Court in this case issued five separate opinions and thus the exact impact of *James Beam* is difficult to

[116] *American Trucking Associations v Smith* 496 US 167 (1990).

[117] *American Trucking Associations v Scheiner* 483 US 266 (1987).

[118] *American Trucking Associations v Smith* (n 116) 178.

[119] Judgment of O'Connor J, joined by Rehnquist CJ, White and Kennedy JJ.

[120] Judgment of Stevens J, joined by Brennan, Marshall, and Blackmun JJ.

[121] *American Trucking Associations v Smith* (n 116) 205–06.

[122] ibid 201.

[123] *James B Beam Distilling v Georgia* 501 US 529 (1991).

[124] *Bacchus Imports v Dias* 468 US 263 (1984).

gauge. Justices Souter and Stevens were of opinion that once the Court applied a rule to litigants in one case, it would have to do so to all similarly situated litigants. Justice White, being of the opinion that 'pure' prospectivity was a settled matter in the precedent of the court,[125] thought that the benefit of *Bacchus* could be extended to the petitioner under any of several suppositions, but a general principle of retroactivity was not one of them. Justice Blackmun (joined by Marshall and Scalia JJ) concluded that the nature of judicial review itself required the retroactive application of new rules. O'Connor J (joined by Rehnquist CJ and Kennedy J) dissented.

On aggregate, there were five judicial votes total[126] for the proposition that court decisions should have retrospective effect and be applicable to any non-final cases still under direct review in the courts.[127] However, it would be difficult to say that *James Beam* on its own was sufficiently clear to fully dispel the spectre of prospective-only effect.

The reversal of the *Chevron Oil* position was made complete with *Harper v Virginia Department of Taxation*. This decision is discussed below under the section addressing the modern United States position on the retroactivity of federal civil decisions.

3.5. THE MODERN APPROACH TO UNCONSTITUTIONALITY AND RETROACTIVITY

3.5.1. Retreating from the *Marbury/Norton* View

As mentioned above, the strong initial stance regarding unconstitutional laws in *Marbury* and *Norton* underwent some significant attenuation after the turn of the 20th century. *Chicot County Drainage v Baxter State Bank*,[128] for example, sees the

[125] For this proposition he relied on: *Cipriano v City of Houma* 395 US 701 (1969).

[126] Souter, Stevens, Blackmun, Marshall and Scalia JJ.

[127] *James B Beam Distilling v Georgia* (n 123) 543–44 (Souter J, Stevens J concurring), 548–49 (Scalia J, Marshall and Blackmun JJ concurring).

[128] *Chicot County Drainage District v Baxter State Bank* 308 US 371 (1940).

Court make significant concessions with respect to the limits of absolute retrospectivity for unconstitutionality. The issue in *Chicot* was a debt-modification arrangement under statute that had been applied for by Chicot County Drainage. As part of that arrangement, certain old obligations were cancelled unless they were presented to the District Court within one year. Although the plaintiff had notice of this, they did not present their bonds to the Court and thus the order made by the Court took no account of them. Subsequently, the statute authorising the debt adjustment was declared unconstitutional. The dispute on appeal between the parties was ultimately whether the original order of the District Court rendered the issue *res judicata* or whether the old bonds could now be recovered. In other words, whether the unconstitutionality of the statute negated judicial orders that were made under it.

The Supreme Court reversed the judgment of the lower court and allowed the plaintiff to recover the bonds. The Supreme Court preferred the view that since the unconstitutionality was declared after it was made, the order should still be allowed to stand. Had Baxter State Bank wished to avail of the unconstitutionality, it would have to have made that contention itself at the original hearing. The Court further held that 'an all-inclusive statement of a principle of absolute retroactive invalidity cannot be justified'.[129] This is difficult to square with the void *ab initio* inclinations of *Norton*. The court in *Chicot* took a dim view of the zealous approach in cases like *Norton*, saying that 'such broad statements as to the effect of a determination of unconstitutionality must be taken with qualifications. The actual existence of a statute, prior to such a determination, is an operative fact, and may have consequences which cannot justly be ignored. The past cannot always be erased by a new judicial declaration'.[130]

An even more decisive statement against the *Norton* view was subsequently made by the Supreme Court in *Lemon v Kurtzman*, wherein Chief Justice Burger wrote that: 'However appealing the logic of *Norton* may have been in the abstract, its aban-

[129] ibid 374.

[130] ibid 374.

donment reflected our recognition that statutory or even judge-made rules of law are hard facts on which people must rely in making decisions and in shaping their conduct'.[131] The abandonment of *Norton* as a statement of law would, therefore, seem almost complete. It would not be true, however, to say that it has entirely vanished from modern American legal thought. Consider this modern statement of the rules regarding unconstitutional statutes:

> The general rule is that an unconstitutional statute, whether federal or state, though having the form and name of law, is in reality no law, but is wholly void, and ineffective for any purpose. Since unconstitutionality dates from the time of its enactment, and not merely from the date of the decision so branding it, an unconstitutional law, in legal contemplation, is as inoperative as if it had never been passed and never existed, that is, it is void *ab initio*. Such a statute leaves the question that it purports to settle just as it would be had the statute not been enacted.[132]

The rhetoric around this issue is apt to confuse, and there seems to be some inconsistency with respect to what accurately states the law. I noted earlier in chapter 2 that the more current view in modern judgments on the topic is that judges can only modify the applicability, rather than the validity, of unconstitutional laws.[133] It is also the only view that is capable of grounding some aspects of older federal law, such as the revival of unconstitutional law (discussed below); that such revivals occurred against the backdrop of the apotheosis of the Supreme Court's invalidity statements gives all the more force to the suggestion that invalidity was always a circumscribed idea in US judicial review. It also forms the view that sits most comfortably with the judicial origins of the power of legislative review in *Marbury* and thus is the better fit with the US constitutional schema generally. As such, I suggest it is the more descriptively accurate view in the US context.

---

[131] *Lemon v Kurtzman* 411 US 192, 199 (1973).

[132] *American Jurisprudence* (5th edn) § 203.

[133] This view is also frequently supported in state courts: *State v Hodge* 128 St3d 1 (Ohio 2010); *Davis v Moore* 772 A2d 204 (DC 2001); *Goodyear Tire and Rubber Company v Vinson* 749 So2d 393 (Ala 1999); *State Ex Rel Moore v Molpus* 578 So2d 624 (Ala 1991).

3.5.2. The New Rules on Retrospectivity

This section clarifies the modern rules on the retrospective application of findings of unconstitutionality in the federal courts. As shall be seen, the *Linkletter* prospectivity standard has been eroded in effectively all areas of law. It has been replaced in each instance by a rule that, generally, favours retrospective application of newly-announced rules.

3.5.2.1. Criminal Cases on Direct Review

The modern rule in the United States for the retrospective application of criminal law to cases taken under direct review stems from *Griffith v Kentucky*.[134] In this case, the court was asked to consider the retrospective effect of the rule in *Batson v Kentucky*,[135] which established that the use of peremptory challenges to strike jurors of the defendant's race from the jury could constitute *prima facie* racial discrimination under the Fourteenth Amendment. The Court concluded that the rule in *Batson* was applicable to litigation pending on direct state or federal review.[136]

The rule for retrospective application stated by the Court is as follows: 'a new rule for the conduct of criminal prosecutions is to be applied retroactively to all cases, state or federal, pending on direct review or not yet final', even where 'the new rule constitutes a "clear break" with the past.'[137] This is a deliberate departure from the earlier 'clear break' exception under *Linkletter* and signalled that confusing categories of this sort were now irrelevant. However, the Court was not blind to the rationale for the 'clear break' exception, and rightly observed that it might infringe on the earlier *Stovall* factors (specifically, the reliance and effect factors), which were still a useful metric in deciding whether final convictions should receive the benefit of a new rule. Notwithstanding this observation, the Court concluded that this was

[134] *Griffith v Kentucky* (n 96).

[135] *Batson v Kentucky* 476 US 79 (1986).

[136] *Griffith v Kentucky* (n 96) 316.

[137] ibid 328.

'precisely the type of case specific analysis that Justice Harlan rejected as inappropriate for cases pending on direct review'.[138] It would also fly in the face of the earlier criticisms of *Linkletter*, which had lamented a rule that did not treat all similarly situated defendants the same.

Thus, the modern standard for retroactivity in criminal cases is a general rule of retroactivity. This rule continues to see modern application.[139]

3.5.2.2. Habeas Corpus Applications

The case of *Teague v Lane*[140] marked a new approach to retrospectivity on the part of the US Supreme Court in *habeas corpus* applications. In effect, 'new' constitutional law cannot be created in, or applied to, federal *habeas corpus* cases. 'New' constitutional law here refers to law that was not 'dictated by precedent existing at the time the petitioner's conviction became final'[141] and which 'breaks new ground or imposes a new obligation on the States or the Federal government'. The Court justified this rule by observing that 'it is not the purpose of the new rule whose benefit the [defendant] seeks, but instead the purposes for which the writ of *habeas corpus* is made available'.[142]

Two narrow exceptions were identified in dissent by Brennan J, which were later approved by a majority of the Court:

> Any time a federal habeas petitioner's claim … would result in the announcement of a new rule of law, … it may only be adjudicated if that rule would [1] plac[e] certain kinds of primary, private individual conduct beyond the power of the criminal law-making authority to proscribe, or [2] if it would mandate new procedures without which the likelihood of an accurate conviction is seriously diminished.[143]

---

[138] *Griffith v Kentucky* (n 96) 327.

[139] See, for example: *Ring v Arizona* 536 US 584 (2002), concerning the retroactive application of *Apprendi v New Jersey* 530 US 466 (2000).

[140] *Teague v Lane* (n 97); aff'd and applied in *Penry v Lynaugh* 492 US 302, 313 (1989).

[141] *Teague v Lane* (n 97) 301.

[142] *Teague v Lane* (n 97) 306. The court was here endorsing Harlan J's dissent in *Mackey v United States* (n 83) (discussed above).

[143] *Teague v Lane* (n 97) 330. This dissent was approved by a majority of the court in *Penry v Lynaugh* 492 US 302 (1989).

A slight gloss has been put on *Teague* more recently by *Greene v Fisher*.[144] This case concerned a rule regarding *habeas corpus* applications taken under the Anti-terrorism and Effective Death Penalty Act 1996 (AEDPA). Under the AEDPA, a federal court cannot grant *habeas corpus* relief to a state prisoner 'with respect to any claim that has been adjudicated on the merits in State court proceedings unless the [state-court adjudication] … resulted in a decision that was contrary to … clearly established Federal law'.[145] The dispute in *Greene* was around the 'clearly established Federal law' aspect of this provision. Greene wished to draw an analogy between *Teague* and the AEDPA, suggesting that because finality[146] was the benchmark for the application of 'new' constitutional rules for the former, it should also work identically for the latter.

Scalia J explained that the AEDPA did not 'codify *Teague*'.[147] They are two separate systems. What counts as 'clearly established Federal law' is the law at the time of the state court trial on the merits; *not* the law at the time the conviction becomes final. The holding in *Greene* thus establishes an exception of sorts to *Teague* under the AEDPA. Under *Teague*, subject to the two exceptions above, 'new' rules cannot avail an applicant on collateral review. Under *Greene*, 'new' rules may be availed of by the applicant on post-conviction review, provided those rules pre-dated the last state review on the merits of the case.

3.5.2.3. Federal Civil Disputes

Judgments possessing only prospective effect were effectively eliminated in civil law by *Harper v Virginia Department of Taxation*.[148] The case has been construed in sub-

[144] *Greene v Fisher* 132 S Ct 38 (2011).

[145] ibid 42–43, citing 28 USC § 2254(d)(1) (2006).

[146] Which occurs when direct state appeals are exhausted and a petition for writ of *certiorari* from the Supreme Court has become time barred or has been disposed of: *Griffith v Kentucky* (n 96) 321, fn 6.

[147] Citing *Horn v Banks* 536 US 266, 272 (2002).

[148] *Harper v Virginia Department of Taxation* (n 102).

sequent court treatments as establishing a 'firm rule of retroactivity'.[149] In *Harper* the Supreme Court was called upon to consider whether its decision in *Davis v Michigan Department of Treasury*[150] (invalidating a Michigan tax on retirement benefits paid by the federal government that exempted retirement benefits paid by the state, on the ground that it violated the constitutional doctrine of intergovernmental tax immunity) had retroactive effect. The issue in *Harper* was whether tax could be claimed back under a similar Virginia statute (that had since been amended) that had breached the principle in *Davis* in cases where the violations had occurred before *Davis* had been decided.

On this occasion, the court gave more definitive guidance than it had previously in *American Trucking v Smith* and *James Beam*. In delivering a single majority opinion it attempted to dispatch lingering ambiguity as to the applicable standard in civil cases:

> When this Court applies a rule of federal law to the parties before it, that rule is the controlling interpretation of federal law and must be given full retroactive effect in all cases still open on direct review and as to all events, regardless of whether such events pre-date or postdate our announcement of the rule.[151]

The analysis of civil law is not so cut-and-dried, however, as the Court did not completely overrule *Chevron Oil*. Rather, the majority opinion was specifically a proscription of what it termed 'selective prospectivity' (application of the rule to the parties of the case, but not other similarly situated parties).[152] It still did not forbid 'pure prospectivity'[153] (the rule does not apply to the parties of the case, nor to any similarly situated party whose claim rests on events pre-dating the decision), which is what it considered the test in *Chevron Oil* provided for. Notwithstanding the continued possibility of pure prospectivity, the general view seems to be that *Harper*

---

[149] *Landgraf v USI Film Producers* (n 113) 279 fn 32.

[150] *Davis v Michigan Department of Treasury* 489 US 803 (1989).

[151] *Harper v Virginia Department of Taxation* (n 102) 97.

[152] '*Harper* overruled *Chevron Oil* insofar as the case (selectively) permitted the prospective-only application of a new rule of law'. *Reynoldsville Casket Co v Hyde* 514 US 749, 752 (1995).

[153] *James B Beam Distilling v Georgia* (n 123) 536.

completed the Supreme Court's turn away from the *Linkletter* standard. Hammer recently summarised the impact of *Harper* as follows:

> With *Harper*, the Supreme Court came virtually full circle in its approach to prospective decisionmaking. The Court has now decisively rejected selective prospectivity in both the criminal and civil contexts and indicated that pure prospectivity may also be forbidden. The two central reasons the Court has repeatedly cited for rejecting prospectivity are: (1) the nature of the judicial function, and (2) the need for the equitable treatment of litigants.[154]

It therefore seems that the retention of pure prospectivity post-*Harper* is purely technical only. Considering the theoretical concerns that grounded the Supreme Court's rejection of selective prospectivity, it seems unlikely that pure prospectivity is likely to be invoked even if it is still technically possible as a matter of precedent.

3.6. AS-APPLIED AND FACIAL CHALLENGES

In each of the other jurisdictions considered in this thesis, all declarations of unconstitutionality are assumed to have *erga omnes* effect. The United States is an exception to this. This section outlines the approach of the United States federal courts, where there is a distinction drawn between legislation being challenged facially or being challenged as-applied. In the case of a facial challenge, the statute is held to be simply unenforceable, with some limited exceptions with respect to state law invalidated in a federal court. In the case of an as-applied challenge, the statute is held to be unenforceable only with respect to a limited set of facts.[155] Unfortunately, this area is rife with confusion and has been fecund ground for academic debate; it is not particularly heartening to have one commentator conclude: '[i]n short, the law is this area is a mess'.[156]

An important point to bear out at the beginning of this discussion is that the distinction between facial and as-applied challenges does not seem to be one of the litig-

[154] Stephen Hammer, 'Retroactivity and Restraint: An Anglo-American Comparison' (2018) 41 Harv J L & Pub Pol'y 409, 422.

[155] Michael Dorf, 'Facial Challenges to State and Federal Statutes' (1994) 46 Stan L Rev 235, 236.

[156] Edward Hartnett, 'Modest Hope for a Modest Roberts Court: Deference, Facial Challenges, and the Comparative Competence of Courts' (2006) 59 SMU L Rev 1735, 1751.

ant's choice; it is not the case that a litigant in the federal courts challenging either a federal or state piece of legislation gets to *opt* between two types of challenge, which have different burdens of proof and different bearings on the ultimate legal status of the impugned statute.[157] Rather, the court will arrive at a determination whether the statute is facially invalid, or simply invalid as-applied (inapplicable) through an application of substantive rules of constitutional law.[158] This is important as it means that the distinction is not one of litigation strategy, and a litigant cannot plead both simultaneously and simply 'hope for the best'. Instead, the court is to decide how to characterise the challenge. Thus, there will be various doctrinal tests in constitutional law that direct a court to consider the statute either facially or as it is applied. Fallon has recently pointed out that most of these tests favour facial invalidity over an as-applied analysis.[159]

The history of this distinction has been uncertain, with older cases seeming to slide between the two different types of challenge without much analysis.[160] As Gorman has observed,[161] some cases saw the Court acting occasionally as though it was

---

[157] This certainly appears to be the view adopted in: *Gonzales v Raich* 545 US 1, 8 (pleading an as-applied claim) 17–20 (court nevertheless adopts a facial analysis) (2005). Cf, however, the dissent of Justice Scalia in *City of Chicago v Morales* 527 US 41, 77–78 (1999).
For academic discussion of this point, supporting the view that it is not a matter of the litigant's choice, see Luke Meier, 'Facial Challenges and Separation of Powers' (2010) 85 Ind L J 1557, 1565–66. It is also worth noting that some have stressed that the parties can opt to frame the litigation in a way that will try to steer the court in the direction of one or the other: Richard Fallon, 'Fact and Fiction about Facial Challenges' (2011) 99 Cal L Rev 914, 947.

[158] Richard Fallon, 'As-Applied and Facial Challenges and Third-Party Standing' (2000) 113 Harv L Rev 1321, 1327–28.

[159] 'A survey of leading cases unmistakably demonstrates that the Court has held statutes wholly invalid under nearly every provision of the Constitution under which it has adjudicated challenges to statutes.' Fallon, 'Fact and Fiction about Facial Challenges' (n 157) 935.

[160] As one commentator put it just over 60 years ago:

> Desire to know and understand the basis upon which the judiciary decides the constitutionality of statutes poses an intriguing problem in analysis and synthesis, for an exhaustive study of it would involve a passage from the strict school of canons of interpretation to the hopelessly loose school of 'corn flakes and judicial disposition.'

Robert Gorman, 'Supreme Court Judgment of State Statute as Unconstitutional on its Face' (1956) 31 Notre Dame L Rev 684, 684.

[161] ibid 689.

examining simply whether the impugned law was facially constitutional;[162] at other times, the Court instead had seemed to prefer to assess whether the law is unconstitutional in light of the specific factual background of the case at bar.[163] Even in more modern cases, the Supreme Court has conceded that the distinction is 'not so well defined that it has some automatic effect or that it must always control the pleadings and disposition in every case involving a constitutional challenge'.[164] There is thus some lingering uncertainty about what analytic criteria (if any) distinguish the two types of challenge.

The first major judicial attempt at clarification as to the facial/as-applied distinction arose in Chief Justice Rehnquist's judgment in *United States v Salerno*.[165] The test laid down by the judge in that case was that a facial challenge to a statue must fail if the statute has *any* constitutional application.[166] Dorf has observed that this brings about a situation where one can 'prevail on a facial challenge only if [one] can also prevail on an as-applied challenge, and even then [one] may lose the facial challenge. Under *Salerno*, a litigant bringing a facial rather than an as-applied challenge gains nothing'.[167] Of course, this misleadingly suggests that the type of challenge is one of the litigant's choice; as mentioned above, this is not the case. The point that *Salerno* lays down a high bar for facial challenges is, however, more problematic. Per-

[162] 'We think that the ordinance is invalid on its face'. *Lovell v City of Griffin* 303 US 444, 451 (1938); 'If on its face the challenged provision is repugnant to the due process clause, specification of details of the offense intended to be charged would not serve to validate it . . . It is the statute, not the accusation under it, that prescribes the rule to govern conduct and warns against transgression'. *Lanzetta v New Jersey* 306 US 451, 453 (1939).

[163] 'But this court must deal with the case in hand and not with imaginary ones. It suffices, therefore, to hold that, as applied to cases like the present, the statute is valid'. *Yazoo & Mississippi Valley Railroad v Jackson Vinegar* 226 US 217, 219–20 (1912); 'Usually, however, the only proper approach takes into consideration both the facts of the case and the construction which the state has placed on the challenged law . . . And in the absence of facts in the light of which the statute may be construed, we have said the proper procedure is not to pass on whether it conflicts with First Amendment rights'. *Kunz v New York* 340 US 290, 304 (1951).

[164] *Citizens United v Federal Election Commission* 130 S Ct 876, 893 (2010).

[165] *United States v Salerno* 481 US 739 (1987).

[166] ibid 745.

[167] Dorf, 'Facial Challenges to State and Federal Statutes' (n 155) 239.

haps because of this exacting standard, there has been much academic commentary exploring the extent to which the *Salerno* doctrine reflects the standard applied by the court in practice.[168]

The *Salerno* standard has been harshly criticised by both the judiciary and the academy. Justice Stevens in particular voiced consistent dissent on the standard, questioning not only whether the approach in *Salerno* is wise in principle,[169] but also whether the court has actually ever really committed itself to the 'no set of circumstances' test at all.[170] On other occasions Justice Stevens simply asserted that *Salerno* is not the proper standard for facial challenges at all.[171] By contrast, the *Salerno* standard found favour with Justice Scalia, who expressed some derision for his colleague's dismissal of the standard on the basis of Dorf's criticisms.[172]

The Court made some attempt to bring more clarity to the proper standard to be applied in *Sabri v United States*.[173] It stressed its general reluctance to entertain facial challenges, suggesting that they would be best used infrequently.[174] The Court rooted its scepticism in the limited ability of courts to scrutinise statutes fairly and holistically given their institutional limitations.[175] It could only consider the statute in the confines of one limited case and only on the basis of the evidence adduced before it.[176] This could cause it to pronounce unnecessarily on constitutional issues

---

[168] Mark Isserles, 'Overcoming Overbreadth: Facial Challenges and the Valid Rule Requirement' (1998) 48 Am U L Rev 359; Scott Keller and Misha Tesytlin, 'Applying Constitutional Decision Rules Versus Invalidating Statutes *in Toto*' (2012) 98 Va L Rev 301.

[169] *Reno v Flores* 507 US 292, 343 fn 28 (1993).

[170] *Janklow v Planned Parenthood* 517 US 1174, 1175 (1995); *Washington v Glucksberg* 521 US 702, 739–40 (1997).

[171] *City of Chicago v Morales* (n 157) 55 fn 22.

[172] 'Justice Stevens asserts that … contrary to the repeated statement of our cases, [the rule in *Salerno*] *never existed*. For that head snapping proposition, he relies upon no less weighty authority than a law review article by Michael C. Dorf.' *Janklow v Planned Parenthood* 517 US 1174, 1180. Emphasis original (1995).

[173] *Sabri v United States* 541 US 600 (2004).

[174] ibid 608.

[175] The idea that judges are inapt to decide 'polycentric' issues was most famously outlined by Lon Fuller in: Lon Fuller, 'The Forms and Limits of Adjudication' (1978) 92 Harv L Rev 353.

[176] *Sabri v United States* (n 173) 609.

and thus fly in the face of the doctrine of avoidance.[177]

More recently, several cases challenging state abortion laws have given the Court further opportunity to reflect on this aspect of its jurisprudence. In *Ayotte v Planned Parenthood of Northern New England*,[178] the Court was called upon to review a decision that held the New Hampshire Parental Notification Prior to Abortion Act, which required minors to notify their parents if they sought to procure an abortion, facially unconstitutional. Although the Supreme Court was sympathetic to the reasoning of the lower court in supporting facial invalidity, even noting that the Court had itself held an abortion statute facially invalid for the self-same reason,[179] it nevertheless castigated the lower court for not considering 'relief more finely drawn'.[180] The Supreme Court preferred the view that in this case partial invalidation (applying only to emergency cases) was the preferable remedy.[181]

The next year, another difficulty with an abortion statute (this time a federal measure) arose in *Gonzales v Carhart*.[182] The Partial-Birth Abortion Ban Act of 2003[183] proscribed a specific technique known as 'partial-birth'. The provision was found unconstitutional in the District Court partly on the basis that it did not provide an exception for emergency cases where this technique would have to be employed to guarantee the health of the mother.[184] The Supreme Court relied on the finding in the Court of Appeal that 'substantial disagreement exists in the medical community regarding whether the procedures prohibited by the Act are ever necessary to preserve a woman's health'[185] in determining that the Act should not be held fa-

[177] The Court was here channelling *United States v Raines* 362 US 17, 22 (1960).

[178] *Ayotte v Planned Parenthood of Northern New England* 546 US 320 (2006).

[179] '[W]e, too, have previously invalidated an abortion statute in its entirety because of the same constitutional flaw'. ibid 330–31.

[180] ibid 331.

[181] ibid 331.

[182] *Gonzales v Carhart* 127 S Ct 1610 (2007).

[183] 18 U.S.C. § 1531

[184] *Gonzales v Carhart* (n 182) 1625.

[185] ibid 1625.

cially invalid. Justice Kennedy took the opportunity to reassert the orthodox view that '[a]s-applied challenges are the basic building blocks of constitutional adjudication'.[186] The Court has thus been quite consistent in stating that as-applied challenges are the norm and facial challenges the rare exception. However, there are more examples of facial invalidation in the cases surveyed than there are examples of as-applied unconstitutionality.[187]

Why the favour for facial challenges? Many scholars have suggested that severability is the central notion in understanding the Supreme Court's use of facial challenges[188] (though some have also contested this claim).[189] Metzger has put the point forcefully, stating that 'existing scholarship generally agrees that the debate regarding the availability of facial challenges is, at bottom, fundamentally a debate about severability'.[190] The question, in the view of these scholars, is whether the unconstitutional applications of a statute can properly be severed from the constitutional applications. The stricture of the *Salerno* doctrine bolsters this view: it is often not difficult to envisage at least *one* constitutional application for a statute. It thus seems attractive to re-conceptualise the issue as one asking whether a statute can really be constitutionally 'pruned' or not. Scholars have theorised the link between severability and facial challenges in various ways.[191] The most prominent approach stems from what has been termed the 'valid rule requirement'—that is, that one always

[186] *Gonzales v Carhart* (n 182) 1639. Citing: Fallon, 'As-Applied and Facial Challenges and Third-Party Standing' (n 158) 1328.

[187] See: *United States v Stevens* 559 US 460 (2010); *Stenberg v Carhart* 530 US 914 (2000); *Planned Parenthood of Southeastern Pennsylvania v Casey* 505 US 833 (1992).

[188] Henry Monaghan, 'Overbreadth' [1981] S Ct Rev 1, 3–6; Dorf, 'Facial Challenges to State and Federal Statutes' (n 155) 249–51; Fallon, 'Fact and Fiction about Facial Challenges' (n 157) 953.

[189] Mark Isserles, 'Overcoming Overbreadth: Facial Challenges and the Valid Rule Requirement' (1998) 48 Am U L Rev 359, 387; Meier (n 157) 1572.

[190] Gillian Metzger, 'Facial Challenges and Federalism' (2005) 105 Colum L Rev 873, 887.

[191] The link between the two is close. Severability is centred around the court removing words from a statute. The as-applied challenge is about removing applications from a statute. Given that a statute's remit of applications will be determined by the words expressed in its provisions, there is a clear link between severability and the as-applied challenge.

has the right to be judged according only to legally valid rules.[192] *Salerno*, then, is taken as establishing an 'irrebuttable presumption that a statute's unconstitutional applications are severable from its constitutional ones'.[193] Even those who have disagreed with the analysis supporting the valid rule requirement[194] still concede some role to severability in the analysis.[195]

The divide between the two types of challenge has come under some fire from some critics. The principal differentiating characteristic separating the two types of challenge is the effects that they can produce: facial challenges simply invalidate legislation, whereas as-applied challenges only make it inapplicable in certain circumstances.[196] However, due to the manner in which judicial review operates in the US federal courts, it is not quite so simple as this. As Fallon has observed:

> When a court rules that a statute is invalid—whether as applied, in part, or on its face—the legal force of its decision resides in doctrines of claim and issue preclusion and of precedent. Under these doctrines, it generally makes no difference whether a court has 'held' that a statute is facially invalid or merely has so reasoned in the course of adjudicating an as-applied challenge.[197]

Despite the considerable criticism with respect to the analytic tenability of the distinction, it is interesting to note that scholars are divided on the issue of whether, in lieu of the distinction, all constitutional challenges should be classified as facial,[198]

---

[192] Henry Monaghan, 'Overbreadth' [1981] S Ct Rev 1, 3; Dorf, 'Facial Challenges to State and Federal Statutes' (n 155) 242–44; Mark Isserles, 'Overcoming Overbreadth: Facial Challenges and the Valid Rule Requirement' (1998) 48 Am U L Rev 359, 389–95; Fallon, 'As-Applied and Facial Challenges and Third-Party Standing' (n 158) 1331–33.

[193] Dorf, 'Facial Challenges to State and Federal Statutes' (n 155) 238.

[194] Matthew Adler, 'Rights, Rules, and the Structure of Constitutional Adjudication: A Response to Professor Fallon' (2000) 113 Harv L Rev 1371, 1395–1406.

[195] Matthew Adler, 'Rights Against Rules: The Moral Structure of American Constitutional Law' (1998) 97 Mich L Rev 1, 158.

[196] Metzger (n 190) 880.

[197] Fallon, 'As-Applied and Facial Challenges and Third-Party Standing' (n 158) 1339.

[198] Adler, 'Rights Against Rules: The Moral Structure of American Constitutional Law' (n 195) 124–32, 157.

as-applied,[199] or both.[200] Things are somewhat clearer when it comes to answering the question of what the court tends to do as a matter of practice. Analysis that has been conducted on whether, as a matter of empirical fact, the US Supreme Court tends to favour as-applied or facial challenges has overwhelmingly suggested that the Court relies on declarations of facial invalidity much more than its rhetoric would often suggest.[201]

3.7. REVIVAL OF UNCONSTITUTIONAL STATUTES

The majority of declarations of unconstitutionality being facial leads neatly into a consideration of the next point: the revival of unconstitutional statutes. Statutes can only really be 'revived' if they were stricken wholesale. Facial declarations are often accompanied by language more severe than that attached to as-applied challenges. Mitchell has recently noted[202] that some judicial opinions in the Supreme Court have held unconstitutional statutes to be void,[203] stricken,[204] or simply not law at all.[205] Mitchell has characterised these statements as committing what he terms the 'writ-of-erasure' fallacy. This fallacy describes a conflict between older authorities that expressly maintain that the United States federal courts have no power to erase or invalidate law[206] and the authorities discussed earlier that suggest that unconstitutional statutes are invalid, struck down, or void. The modern expression of this

---

[199] Fallon, 'As-Applied and Facial Challenges and Third-Party Standing' (n 158) 1335–41. It is notable that Fallon seems to have recanted this view in more recent work: Fallon, 'Fact and Fiction about Facial Challenges' (n 157).

[200] David Franklin, 'Looking Through Both Ends of the Telescope: Facial Challenges and the Roberts Court' (2009) 36 Hastings Const L Q 689, 689–90.

[201] For a detailed empirical analysis, see: Fallon, 'Fact and Fiction about Facial Challenges' (n 157) 940–42.

[202] Mitchell (n 24).

[203] *Marbury* (n 3) 177; *Free Enterprise Fund v Public Co Accounting Oversight Board* (n 14) 510.

[204] *Citizens United* (n 164) 346; *Harris v Arizona Independent Redistricting Commission* 578 US — (2016).

[205] *Chicago, Indianapolis & Louisville Railway Company v Hackett* 228 US 559 (1913); *Norton* (n 26); *Reynoldsville Casket Company v Hyde* 514 US 749 (1995).

[206] *Perez v Ledesma* 401 US 82, 124 (1971); *Steffel v Thompson* 415 US 452, 469 (1974).

tension is in the facial and as-applied distinction discussed above.

It is worth noting that this divergence in views can have real consequences. A phenomenon that might be termed 'zombie law' has arisen on rare occasions because of the 'suspension of application' view; that is, occasionally unconstitutional statutes have had new life breathed into them as a result of a reversal of fortunes in the Supreme Court.[207] There have been a few examples of this that I will describe below.

The most well-known example of this occurred when, in 1923, the United States Supreme Court held in *Adkins v Children's Hospital*[208] that a federal minimum wage law for women was unconstitutional. This decision was later overruled in *West Coast Hotel Co v Parrish*, decided in 1937.[209] This gave rise to the question of the validity of the original stricken minimum wage law. Attorney General Cummings advised that the statute had only had its effectiveness suspended by the 1923 holding, and thus after the 1937 decision it was again wholly effective. He maintained that 'the courts have no power to repeal or abolish a statute, and that notwithstanding a decision holding it unconstitutional, a statute continues to remain on the statute books.'[210] As described in chapter 2, in the context of the basis of judicial review in the United States federal courts, the federal courts seem to understand their practice as enjoining official action grounded in unconstitutional law, rather than striking that unconstitutional law down.

A less well-cited example of a similar phenomenon occurred earlier in *In Re Rahrer*.[211] The petitioner in this case had breached the state of Kansas's prohibitory law on the sale of alcohol. The Supreme Court had previously, in *Leisy v Hardin*,[212]

---

[207] The most in-depth academic treatment of this is William Treanor and Gene Sperling, 'Prospective Overruling and the Revival of "Unconstitutional" Statutes' (1993) 93 Colum L Rev 1902.

[208] *Adkins v Children's Hospital* 561 US 525 (1923).

[209] *West Coast Hotel Co v Parrish* 300 US 379 (1937).

[210] 39 Ops Atty Gen 22 (1937)

[211] *In Re Rahrer* 140 US 545 (1891).

[212] *Leisy v Hardin* 135 US 100 (1890).

declared it unconstitutional for a state statutory instrument to purport to regulate the sale and distribution of goods legally imported into that state; this was the power of Congress under the Commerce Clause. In between the decision in *Leisy* and *Re Rahrer*, Congress enacted what was known as the Wilson Bill, which empowered states to regulate imported intoxicated liquors. For present purposes, what is most noteworthy is that the state of Kansas did not have to re-enact its statutory prohibition notwithstanding its unconstitutionality under *Leisy* following its restoration to constitutionality under the Wilson Bill. The interjection of Congress was sufficient to resuscitate the Kansas statute as the Wilson Bill effectively overruled the Supreme Court in *Leisy*. Fuller CJ, delivering judgment in *Re Rahrer*, did not think that the statutes impugned by *Leisy* had been annulled but rather 'limited [in their] operation to property strictly within the jurisdiction of the state'.[213] Thus, the Wilson Bill was not acting retrospectively to impose liability on the petitioner in *Re Rahrer*; rather, it was simply adjusting the jurisdictional ambit of the still-extant state law.

Another example, though also a dated and somewhat extreme one, is to be found in the *Legal Tender Cases*[214] that appeared before the United States Supreme Court concerning the constitutionality of paper money not redeemable in gold or silver (this occurred against the backdrop of the civil war, with the issuance of paper money being used to fund the war). In *Hepburn v Griswold*[215] the Supreme Court held that the Legal Tender Act of 1862 constituted a deprivation of the right to property under the Fifth Amendment as it mandated the acceptance of greenbacks in satisfaction of any debt. Shortly afterwards, *Hepburn* was overruled in the *Legal Tender Cases* (with two new appointments to the Supreme Court nominated by President Grant in the interim). The Court did not specifically attend to the issue of whether the relevant provisions of the Legal Tender Act would need to be re-passed; instead, it seemed to assume as a matter of course that the Act would be enforceable after

[213] *In Re Rahrer* (n 211) 563.

[214] *Legal Tender Cases* 79 US 457 (1870).

[215] *Hepburn v Griswold* 75 US 603 (1868).

Hepburn was overruled.

The majority position among academics who have analysed this aspect of American law,[216] with some dissent,[217] is that the analysis of Advocate General Cummings was correct and a law may be rendered effective again if the precedent that 'blocked' it is overruled. More than a century ago, the New Jersey Court of Errors and Appeals offered the following prescient summary of what has now become the orthodox position in the federal courts:

> [The judicial role], with respect to legislation deemed unconstitutional, is not exercised *in rem*, but always *in personam*. The Supreme Court … simply ignores statutes deemed unconstitutional. … An unconstitutional statute is not merely blank paper. The solemn act of the legislature is a fact to be reckoned with. Nowhere has power been vested to expunge it or remove it from its proper place among statutes.[218]

This seems to confirm the revival of 'dead' laws as a possibility if the Supreme Court no longer has reason to 'ignore' the statute deemed unconstitutional. This could happen where, as in the *Legal Tender Cases* there is a revision of the correct constitutional standard to be applied.

3.8. CONCLUSION

One of the difficulties in accounting for the United States' practice of unconstitutionality is the extent to which it has oscillated over its long lifetime. As seen above, the Supreme Court has experimented with various retrospectivity and prospectivity doctrines, and the distinction between facial and as-applied challenges suggests that whatever notion of 'validity' the court is using may be less radical than statements to

[216] Stuart Buck and Mark Rienzi, 'Federal Courts, Overbreadth, and Vagueness: Guiding Principles for Constitutional Challenges to Uninterpreted State Statutes' [2002] Utah L Rev 381, 425; Richard Fallon, 'Making Sense of Overbreadth' (1991) 100 Yale L J 876, 876; Melville Nimmer, 'A Proposal for Judicial Validation of a Previously Unconstitutional Law: The Civil Rights Act of 1875' (1965) 65 Colum L Rev 1394, 1398.

[217] Treanor and Sperling (n 207).

[218] *Allison v Corker* 67 NJL 596, 601 (1902).

the effect that unconstitutional laws are 'void'[219] or 'struck down'[220] might suggest. As Shapiro put it: 'No matter what language is used in a judicial opinion, a federal court cannot repeal a duly enacted statute of any legislative authority'.[221]

It is surprising that the fundamental effect of a finding of unconstitutionality could remain a somewhat unsettled question after such a long-standing practice of judicial review, but this seems to be the case. I do not advance a view on what the better answer to this question should be from the perspective of an internal participant of the United States legal system. For the purposes of a comparative study, what is interesting is that disagreement can occur over this question even within a jurisdiction, as opposed to between jurisdictions. This possibility of intra-jurisdictional disagreement suggests that a jurisdiction does not necessarily have to maintain one theory of unconstitutionality for all time. Theories of unconstitutionality can differ across time, as well as across space in different legal systems.

Another interesting lesson to draw from the United States' jurisprudence is the difficulties that can arise in tampering with the temporal effects of unconstitutionality. The Warren Court's brief experiment in dabbling with prospective-only effect following unconstitutionality ended in failure. There are theoretical and general problems with prospective judgments and unconstitutionality that are analysed later in this thesis, but the United States' experience is a useful object lesson on the strength of these objections.

Finally, the important (albeit rare) phenomenon of the revival of unconstitutional laws is worth underscoring. In most cases, what seems to occur is that a legal norm is declared unconstitutional in one case, and this revives a *different provision* endorsing the same norm.[222] An exception to this is the *Legal Tender Cases*, where a

---

[219] *Marbury v Madison* 5 US 137, 177 (1803); *Norton v Shelby County* 118 US 425, 426 (1886); *Free Enterprise Fund v Public Co Accounting Oversight Board* 561 US 477, 510 (2010).

[220] *Citizens United v Federal Election Commission* 130 S Ct 876, 346 (2010); *Harris v Arizona Independent Redistricting Commission* 578 US — (2016).

[221] David Shapiro, 'State Courts and Federal Declaratory Judgments' (1979) 74 Nw U L Rev 759, 767.

[222] So, for example, *Leisy* (discussed above) concerned an Iowa law, but the petition in *Re Rahrer* was concerned with a Kansas law. Similarly, *Adkins* concerned federal, District of Columbia law,

single legislative provision was stricken and the judgment constituting the unconstitutionality was itself overruled, restoring the provisions providing for paper money as legal tender.

whereas *Parrish* concerned Washington state law.

4 | Canada

CHAPTER OVERVIEW

4.1. Introduction 101

4.2. Doctrines Avoiding or Curtailing Unconstitutionality 102

4.2.1. The Presumption of Constitutionality 103

4.2.2. Reading In 104

4.2.3. Severance 105

4.3. Early Case Law: Unconstitutionality as Voidness/Invalidity . . . 105

4.3.1. Restitution for Unconstitutional Taxation 106

4.3.2. Review under the Charter: *Big M Drug Mart* 108

4.3.3. The *Manitoba Language Rights* Reference 110

4.4. The Rise of the Suspended Declaration 113

4.4.1. *Schachter v Canada* 113

4.4.2. The Proliferation of Suspended Declarations: *Bedford* and *Carter* 114

4.5. The Revival of Unconstitutional Law 116

4.6. Conclusion 117

4.1. INTRODUCTION

The Canadian Constitution is the second-oldest constitutional instrument discussed in this thesis, originating in the British North America Act 1867/Constitution Act 1867 and ranging up to the Constitution Act 1982. As briefly discussed in chapter 2, although judicial review pre-dated the 1982 Act, it was a more limited power based

on federalism concerns; the federal courts could review provincial laws to examine their compliance with the arrangements for the devolution of legislative competence. The historical basis of the evolution of this power is complex, and I will not rehearse it here.[1] However, it was clear from an early stage that the Supreme Court of Canada, upon its first establishment under the Supreme and Exchequer Courts Act 1875, considered that it had a power of judicial review of legislation.[2]

The power of judicial review was given textual footing for the first time by section 52 of the Constitution Act 1982,[3] which requires that laws that are inconsistent with the provisions of the Constitution are 'of no force or effect'. This requires courts to intervene in such cases and, as such, assumes the existence of a judicial review power.

As with the United States chapter, the discussion here will begin with doctrines that restrain or curtail unconstitutionality. An added nuance is required in this classification, however, that was unnecessary in the context of the United States. The suspension of a declaration of unconstitutionality does not qualify as a limiting or curtailing doctrine as I intend to describe such doctrines. Suspension limits the consequences of unconstitutionality once it has been established, but the point of the limiting doctrines I discuss first is to avoid unconstitutionality *tout court* where possible. They are preventative measures against constitutionality, rather than the emergency surgery entailed by a suspended declaration. Suspended declarations are, therefore, treated separately.

4.2. DOCTRINES AVOIDING OR CURTAILING UNCONSTITUTIONALITY

The Supreme Court's leading decision on remedies is *Schachter v Canada*.[4] This case is therefore a natural fulcrum for the discussion in this section. As constitutional

[1] See generally: Jennifer Smith, 'The Origins of Judicial Review in Canada' (1983) 16 Canadian Journal of Political Science 115.

[2] *Severn v The Queen* [1878] 2 SCR 70 (SCC); *Valin v Langlois* [1879] 3 SCR 1 (SCC).

[3] Smith (n 1) 134.

[4] *Schachter v Canada* [1992] 2 SCR 679 (SCC).

limitation doctrines are not the chief subject of inquiry of this thesis, a detailed account of each of these doctrines is not supplied here.

4.2.1. The Presumption of Constitutionality

The Canadian courts seem to have rejected a presumption of constitutionality, at least as would apply to the Charter of Rights and Freedoms. There is no presumption that legislation is Charter-compliant. The earliest statement of the Supreme Court to this effect was in *Attorney General of Manitoba v Metropolitan Stores (MTS) Ltd*.[5] However, some caution is necessary in delineating exactly what the Court rejected. Beetz J outlined his understanding of 'the presumption of constitutionality' as requiring: 'that a legislative provision challenged on the basis of the Charter must be presumed to be consistent with the Charter and of full force and effect'.[6]

It was this definition that the Court rejected, maintaining that 'the presumption of constitutional validity . . . whether it is applied to laws enacted prior to the Charter or after the Charter, is not compatible with the innovative and evolutive character of this constitutional instrument'.[7] However, the Court made equally clear that it was not opposing the general rule that the onus of proof lies with the party making a particular claim.[8] Nor was it opposing a canon of construction whereby statutes are to be interpreted compatibly with the Charter where such readings are possible.[9]

This rejection of the presumption has been supported in subsequent cases.[10] There is therefore no presumption of constitutionality in Charter cases.

---

5 *Manitoba (AG) v Metropolitan Stores Ltd* 1987 CanLII 79 (SCC), [1987] 1 SCR 110.

6 ibid [14].

7 ibid [23].

8 ibid [25].

9 ibid [26].

10 *R v Zundel* 1992 CanLII 75 (SCC), [1992] 2 SCR 731, 758; *Harper v Canada (Attorney General)* 2000 SCC 57 (CanLII), [2000] 2 SCR 764 [33]dissent of Major J.

4.2.2. Reading In

The Supreme Court of Canada endorsed reading in in the *Schachter* case.[11] This remedy was, in the view of the court, a logical corollary of severance:

> [E]xtension by way of reading in is closely akin to the practice of severance. The difference is the manner in which the extent of the inconsistency is defined. In the usual case of severance the inconsistency is defined as something improperly included in the statute which can be severed and struck down. In the case of reading in the inconsistency is defined as what the statute wrongly excludes rather than what it wrongly includes. Where the inconsistency is defined as what the statute excludes, the logical result of declaring inoperative that inconsistency may be to include the excluded group within the statutory scheme.[12]

Similar to the justification offered above for the adoption of reading in in US equality cases, the Supreme Court of Canada justified reading in on the basis that it shows greater respect for the role of the legislature. It also added the general justification that reading in would occasionally be required to respect the purposes of the Charter as a whole.[13] In determining whether severance or reading in would be appropriate, Lamer CJ identified three criteria.[14] Only if these three criteria are met can severance or reading in be granted:

> A. the legislative objective is obvious … and severance or reading in would further that objective, or constitute a lesser interference with that objective than would striking down;
>
> B. the choice of means used by the legislature to further that objective is not so unequivocal that severance/reading in would constitute an unacceptable intrusion into the legislative domain; and,
>
> C. severance or reading in would not involve an intrusion into legislative budgetary decisions so substantial as to change the nature of the legislative scheme in question.[15]

The test, therefore, is predicated on the intent of the legislature (A), faithfulness to that intent on the part of the court (B) and a principle of minimal intrusion (C).

---

[11] *Bedford* (n 4).

[12] ibid 698 (Lamer CJ).

[13] ibid 700–02 (Lamer CJ).

[14] ibid 705–15 (Lamer CJ).

[15] ibid 718 (Lamer CJ).

This remedy has seen particularly fruitful application in Canada in marriage equality cases, with protection for common law spouses[16] and gay and lesbian couples[17] having been read into multiple instruments.

4.2.3. Severance

Severance is also practised in Canada. The rationale for severance is very closely linked to that of reading in, discussed above in the context of the *Schachter* case. The Canadian approach to constitutional remedies is, in general, quite homogeneous as many of the remedies revolve around two chief concerns: respect for the role of the legislature and respect for the purposes of the Charter.[18] As mentioned above, the test for the availability of severance is the same as that applied to reading in.

4.3. EARLY CASE LAW: UNCONSTITUTIONALITY AS VOIDNESS/INVALIDITY

This section charts the early case law on unconstitutionality in the Supreme Court of Canada. As shall be seen, significant strands of this case law point to a standard of unconstitutionality entailing voidness *ab initio*. These baseline assumptions are important to establish clearly as although modern Canadian practice has largely been taken over by the suspended declaration, this innovation has not disturbed the understanding of how unconstitutionality operates. At least when it comes to the more nuanced, theoretical points of unconstitutionality, the suspended declaration is more of a sideways dodge than a significant advancement or innovation. These early cases are thus of exceptional importance to clarify before pragmatic remedial considerations are added to the analysis.

---

[16] *Miron v Trudel* [1995] 2 SCR 418 (SCC).

[17] *Vriend v Alberta* [1998] 1 SCR 493 (SCC).

[18] Kent Roach, 'Enforcement of the Charter—Subsections 24(1) and 52(1)' (2013) 62 SCLR (2d) 473, 504.

4.3.1. Restitution for Unconstitutional Taxation

The Supreme Court of Canada, much like the United States Supreme Court in the previous chapter, has been required to consider the retrospectivity of unconstitutional legislation in the context of restitution of unlawful taxation. Historically, the Canadian courts adopted a somewhat confusing position on this point. Where legislation was invalid but constitutional, the courts dealt with the dispute as a quasi-contractual or restitutionary claim.[19] On the constitutional side, the early indications were promising: in *Amax Potash v Government of Saskatchewan* the Supreme Court of Canada held that a statute purporting to bar the recovery of taxes that had been paid under an unconstitutional statute was, itself, unconstitutional.[20] The Supreme Court adopted quite a strident stance on this point:

> The principle governing this appeal can be shortly and simply expressed in these terms: if a statute is found to be *ultra vires* the legislature which enacted it, legislation which would have the effect of attaching legal consequences to acts done pursuant to that invalid law must equally be *ultra vires* because it relates to the same subject-matter as that which was involved in the prior legislation. If a state cannot take by unconstitutional means it cannot retain by unconstitutional means.[21]

The government of Saskatchewan was therefore liable for the taxes paid under the provision that was deemed unconstitutional. As this case only concerned an interlocutory application, it did not itself deal with the constitutionality of potash mining regulations *tout court*.

Subsequently, in *Air Canada v British Columbia*, the Supreme Court carved out a significant exception to the *Amax Potash* case, allowing for the imposition of a new and valid retroactive tax that would allow a legislature to block recovery of unconstitutional taxes.[22] The taxes concerned in *Air Canada* were indirect taxes. These breached the provincial power of taxation, which covered only direct taxes.[23]

---

[19] Peter Hogg, *Constitutional Law of Canada* (5th edn, Carswell 2007) 58-15.

[20] *Amax Potash v Government of Saskatchewan* 1976 CanLII 15 (SCC), [1976] 2 SCR 576.

[21] ibid 592.

[22] *Air Canada v British Columbia* 1989 CanLII 95 (SCC), [1989] 1 SCR 1161.

[23] Direct taxes are those directly paid to the government. Indirect taxes are those paid to the govern-

The British Columbia legislature stopped the levy in 1976, and replaced it with a valid, direct tax. This tax was then amended in 1981 to make it apply retroactively to the period 1974 to 1976. La Forest J considered that there should be a general principle in restitution that recovery for unconstitutional taxes be barred.[24]

This principle was revisited, and dismissed, by the Supreme Court in the *Kingstreet Investments v New Brunswick* case.[25] At issue here was a New Brunswick levy imposing a user charge on night clubs licensed to sell alcoholic beverages. This was found to be an unconstitutional indirect tax. New Brunswick, in its defence, attempted to rely on La Forest J's proposal to bar recovery of unconstitutional taxes. The Court thus had to consider La Forest J's principle that legislatures should (at least by default) be effectively immune from suit for recovery of unconstitutionally collected taxes. Bastarache J rejected this view, on the basis that it would undermine the rule of law for legislatures to be able to collect and retain tax that was obtained *ultra vires*.[26] He was unconvinced by the policy motivations advanced by La Forest J in favour of his immunity rule—that fiscal chaos might otherwise result—as he felt that a suspended declaration could avoid such consequences.[27]

In a more significant break from the past, Bastarache J also declared that recovery for unconstitutional taxes was a matter 'of constitutional right'.[28] Quite apart from the private law of restitution, there was a constitutional law cause of action and remedy available in the case of unconstitutional tax statutes.[29] *Kingstreet* is thus important for the more general reason that it sees the Supreme Court of Canada reassert the orthodoxy that unconstitutional action is *ultra vires*, void, and of no ef-

ment by a third party who must then pay the government; these often take the form of value-added tax, import duties etc.

[24] *Air Canada v British Columbia* (n 22) 1206.

[25] *Kingstreet Investments Ltd v New Brunswick (Finance)* 2007 SCC 1 (CanLII), [2007] 1 SCR 3.

[26] ibid [15].

[27] This being precisely the result achieved in a prior case dealing with an unconstitutional Ontario probate fee: *In Re Eurig Estate* [1998] 2 SCR 656 (SCC).

[28] *Kingstreet Investments Ltd v New Brunswick (Finance)* (n 25) [34].

[29] ibid [40].

fect, though with the usual proviso that, by dint of suspended declarations or other legal doctrines, they may nevertheless be given some limited continued application.

4.3.2. Review under the Charter: *Big M Drug Mart*

The first major case to consider the effect of a declaration of unconstitutionality for review of rights under the Charter of Fundamental Rights and Freedoms occurred in the *Big M Drug Mart* case.[30] At issue in *Big M* was the Lord's Day Act 1970, which forbade the sale of goods on a Sunday. The respondent, Big M Drug Mart, alleged that section 4 of the Lord's Day Act was an infringement on the right to freedom of conscience and religion guaranteed by section 2 of the Charter. The Supreme Court struck the Lord's Day Act 1970 down in its entirety as unconstitutional as it was found to effectively enforce a sectarian Christian ideal. In the course of its judgment, the court gave some hint as to how it conceived of the power in section 52:

> Section 52 sets out the fundamental principle of constitutional law that the Constitution is supreme. The undoubted corollary to be drawn from this principle is that no one can be convicted of an offence under an unconstitutional law. … Any accused, whether corporate or individual, may defend a criminal charge by arguing that the law under which the charge is brought is constitutionally invalid. … Whether a corporation can enjoy or exercise freedom of religion is therefore irrelevant[.] … A law which itself infringes religious freedom is, by that reason alone, inconsistent with s. 2(a) of the Charter and it matters not whether the accused is a Christian, Jew, Muslim, Hindu, Buddhist, atheist, agnostic or whether an individual or a corporation. *It is the nature of the law, not the status of the accused, that is in issue.*[31]

The respondent company therefore did not itself need to possess a Charter right to freedom of conscience and religion. All that mattered was that some plausible conception of this freedom was violated by the Lord's Day Act and that this made the Act invalid. This reasoning might be taken to imply at least two further tacit assumptions: *(i) legislative nullity:* the Constitution sets an autonomous and self-executing standard for legal validity, which does not require judicial application or

[30] *R v Big M Drug Mart Ltd* [1985] 1 SCR 295 (SCC).

[31] ibid [38]–[41] (emphasis added).

cognizance to be activated, and *(ii) legal validity = existence ≠ applicability:* legal validity is zero-sum in the sense that it terminates the invalid law; in other words, it is not a matter of the law being unconstitutional in its *application* to some factual matrices and not others.

Proposition *(i)* seems implied in the way the court envisages that standing is irrelevant to the application of section 52. If standing is irrelevant in this way, it seems to be because constitutional validity is seen as something that is acontextual. This acontextual view of unconstitutionality is also supported by *(ii)*. This presents unconstitutionality as a matter of internal constitutional logic. All this said, the holding in *Big M* does not of itself prove that the Canadian courts think of the constitution as regulating its own validity. Proposition *(i)* is compatible with, but not necessarily entailed by, the statement of the court. It would be equally possible to think that judges change the nature of the law by adjudicating on the validity or invalidity of an act. However, it is worth noting that in the later case of *Nova Scotia v Martin*, the Court expressly endorsed proposition *(i)*, holding that the invalidity of an unconstitutional law 'does not arise from the fact of its having been declared unconstitutional by a court, but from the operation of s 52(1)'.[32] Additionally, although the language of voidness *ab initio* does not appear anywhere in this first consideration of section 52,[33] the court's understanding of 'no force or effect' seems tantamount to voidness.[34]

---

[32] *Nova Scotia (Workers' Compensation Board) v Martin; Nova Scotia (Workers' Compensation Board) v Laseur* [2003] 2 SCR 504 (SCC).

[33] The resurgence of this terminology was likely a reflection of the old Colonial Laws Validity Act 1865 discussed above. The understanding that *ultra vires* acts were to be treated as void *ab initio* is one that goes back at least as far as the late 1800s in Supreme Court of Canada jurisprudence: *McCraken v McIntyre* 1877 CanLII 16 (SCC), [1877] 1 SCR 479, 529; *Ottawa Agriculture Insurance Co v Sheridan* 1880 CanLII 21 (SCC), [1879] 5 SCR 157, 159; *McSorley v The Mayor of the City of St John* 1882 CanLII 31 (SCC), [1882] 6 SCR 531, 533.

[34] These two terms have been used very closely together and do not seem to be treated with much analytical distinction. See, for example (albeit in the context of a court martial): *R v Corporal SA Strong* 2008 CM 3019 (CanLII) [24]–[25]. In a more traditionally judicial context, the Alberta Court of Appeal observed that: 'The direct result of a declaration of unconstitutionality is that the impugned legislation is rendered void *ab initio*. The legislation is, and always has been, invalid and of no force or effect.' *Turigan v Alberta* 1988 ABCA 333, [1988] 6 WWR 673 [3].

Proposition *(ii)* seems to be required on the basis that Big M did not itself need to be protected by the reason for the unconstitutionality (the violation of freedom of religion). All it required was that the law was unconstitutional *somehow*. In other words, the Court did not require the reason for the unconstitutionality to be applicable to Big M's own circumstances. This suggests the Court deals with legal validity and existence as joint concepts that are independent of the applicability of law.[35] In light of these arguments, there are good grounds to consider the Supreme Court of Canada as committed to propositions *(i)* and *(ii)*.

More generally, the acknowledgement by the Court that something regarding the nature of law is at play in these cases throws the jurisprudential undertones of this area of constitutional law into sharp relief and calls for a jurisprudential answer to this question. It is also notable that in this initial treatment of constitutional invalidity the remedy was also considered to be 'declaratory',[36] thus eschewing the notion that the court is constituting the invalidity as opposed to merely recognising it. This is a further hint that the court in *Big M* was at least tacitly relying on something like proposition *(i)* above in its reasoning.

4.3.3. The *Manitoba Language Rights* Reference

If indeed the Court was leaning towards the view that the constitution regulates its own validity in *Big M*—a view that would render every unconstitutional law 'void'—the logical consequences of this view were shortly presented in stark fashion in the *Re Manitoba Language Rights* reference.[37] In 1890, the Manitoba legislature had enacted the Official Language Act, which provided that statutes only needed to be published in English. It was not until almost a hundred years later in the *Forest* case that

[35] It is, of course, true that the criminal prohibition of the sale of goods on Sunday was applicable to Big M. It would not have had a case otherwise. And there is a difference between a court *regulating* the applicability of law by the operation of unconstitutionality, and a court considering that unconstitutionality is *itself* regulated by applicability. The point in *Big M* supporting proposition *(ii)* is thus only suggestive at most. There are other, less transparent, disavowals by the Supreme Court of Canada regarding applicability analysis that are considered below.

[36] *Big M Drug Mart* (n 30) [35].

[37] *In Re Manitoba Language Rights* [1985] 1 SCR 721 (SCC).

this came to light as unconstitutional.[38] The result of the *Forest* decision was that *all* Manitoba statutes needed to be in both English and French, though they had only ever been enacted in the former since 1890 on foot of the Official Language Act. The legality of the statutes was not considered in *Forest*, but it was considered in the later *Bilodeau* case.[39] The government, concerned by the potential impact of the issues raised in *Bilodeau*, requested a reference to the Supreme Court on the constitutionality of the English-only statutes. Specifically, the Court was called to adjudicate on whether the failure to translate Manitoba statutes into French violated section 133 of the Constitution Act 1867 and section 23 of the Manitoba Act 1870 (a federal and state constitutional protection of language rights).

The Court found that the contravention of the bilingualism provisions of the Constitution and Manitoba Acts resulted in clear unconstitutionality. This led to the uncomfortable conclusion that the vast majority of the Manitoba statute book was invalid (pre-1980 statutes, and only some few post-1980 statutes, would have been saved). Indeed, the Manitoba legislature itself would have been invalid, having been changed by several English-only post-1890 laws. According to section 52(1), all these monolingual laws would have to be held to be 'of no force or effect'. As Hogg has observed, this placed Manitoba in a catch-22 situation that would destroy all its existing law, and also remove any possibility of enacting provincial law in Manitoba as it would lack a legislature.[40] The administratively catastrophic effects of this finding were avoided by granting a suspended declaration of invalidity, a remedy that has since become well-established in the court's remedial arsenal.[41] In

[38] *Attorney General of Manitoba v Forest* 1979 CanLII 242 (SCC), [1979] 2 SCR 1032.

[39] *Bilodeau v Attorney General of Manitoba* [1981] 5 WWR 393 (MCOA).

[40] Hogg, *Constitutional Law of Canada* (n 19) 58-21.

[41] *Sinclair v Quebec* [1992] 1 SCR 579 (SCC); *In Re Remuneration of Judges (No 2)* [1998] 1 SCR 4 (SCC); *In Re Eurig Estate* (n 27); *M v H* [1999] 2 SCR 3 (SCC); *UFCW v KMart Canada* [1999] 2 SCR 1083 (SCC); *UFCW v KMart Canada* [1999] 2 SCR 1083 (SCC); *Mackin v New Brunswick* [2002] 1 SCR 405 (SCC); *Figueroa v Canada (Attorney General)* [2003] 1 SCR 912 (SCC); *Trociuk v British Columbia (Attorney General)* [2003] 1 SCR 835 (SCC); *Nova Scotia (Workers' Compensation Board) v Martin; Nova Scotia (Workers' Compensation Board) v Laseur* (n 32); *R v Demers* [2004] 2 SCR 489 (SCC); *Charkaoui v Canada (Citizenship and Immigration)* [2007] 1 SCR 350 (SCC); *Health Services and Support – Facilities Subsector Bargaining Association v British Colom-*

principle, however, the legal norms of the Manitoba statute book were void *ab initio*. Again, the court seemed to reaffirm that the Constitution is self-executing when it comes to section 52: '[The Constitution] is, as s 52 of the Constitution Act 1982 declares, the "supreme law" of the nation … *unsuffering of laws inconsistent with it*'.[42] It might seem odd to simultaneously claim that invalidity is timeless and to impose a definite time frame on a court's order following that invalidity, but that is what was done by the court in the *Manitoba Language Rights* reference.

Similar cases arose in both Saskatchewan and Alberta. In 1877, a time during which half of the population of the Northwest Territories was French speaking, the federal parliament enacted a law requiring legislation to be published in both English and French. Saskatchewan was originally a part of the Northwest Territories. When Saskatchewan was created by the Saskatchewan Act 1905, the French-speaking population was much reduced, and so it was assumed that the language law was irrelevant and inapplicable.[43] *R v Mercure*,[44] then, involved a francophone Saskatchewan resident defending a speeding charge on the basis that the legislation was invalid for being monolingual. On the basis of the *Manitoba Language Rights* reference, it was clear that the Supreme Court could not but find for the applicant.

Alberta, like Saskatchewan, was carved out of the Northwest Territories in 1905 and had similarly enacted statutes only in English, ignoring the 1877 language law. *R v Paquette*[45] then, similarly, compelled the Court to again confront the inescapable conclusion that the Alberta statute book was as void as those in Manitoba and Saskatchewan. However, a very important difference distinguishes the Saskatchewan and Alberta cases: it was not a *constitutional* requirement in either province that statutes

bia [2007] 2 SCR 391 (SCC); *Nguyen v Quebéc (Education, Recreation and Sports)* [2009] 3 SCR 208 (SCC); Sujit Choudhry and Kent Roach, 'Putting the Past Behind Us? Prospective Judicial and Legislative Constitutional Remedies' (2003) 21 Supreme Court Law Review (2nd) 205, 253–54.

[42] *Manitoba Language Rights* (n 37) [48] (emphasis added).

[43] Some mention of the connection between desuetude and applicability is made in chapter 8. Few common law countries accept a formal doctrine of desuetude.

[44] *R v Mercure* [1988] 1 SCR 234 (SCC).

[45] *R v Paquette* [1990] 2 SCR 1103 (SCC).

be enacted bilingually. In both cases it was a result of the lingering 1877 act. Thus, the Court suggested that a simple solution would be for the legislatures of Saskatchewan and Alberta to enact a single bilingual act, which would have the effect of repealing the 1877 law and retrospectively declaring all prior monolingual law valid. Both legislatures adopted this solution.[46]

4.4. THE RISE OF THE SUSPENDED DECLARATION

The Supreme Court of Canada was the first court to suspend a declaration of unconstitutionality, a finding it felt no doubt compelled to make against the background of *Re Manitoba Language Rights*. Since then, however, the remedy has become quite established in the Court's arsenal and has even become more the exception than the norm. This section traces this development.

4.4.1. *Schachter v Canada*

In *Schachter v Canada*, the Supreme Court of Canada's leading statement on constitutional remedies, the Court affirmed its use of suspended declarations.[47] It envisaged that the remedy should turn on considerations of the immediate effect of the declaration on the public and not the respective roles of the courts and legislature,[48] and that such declarations should only be made where immediate invalidity would pose some public danger or 'otherwise threaten[] the rule of law'.[49] This demonstrates how the rule of law has been a keystone concept for the Supreme Court of Canada.[50] Subsequent cases have diluted the *Schachter* standard, and the suspen-

[46] The Language Act, SS 1988, c L-6.1 (Saskatchewan); Languages Act, SA 1988, c L-7.5 (Alberta).

[47] There had been a further award of a suspended declaration between *Re Manitoba Language Rights* and *Schachter* in *R v Swain* 1991 CanLII 104 (SCC), [1991] 1 SCR 933. The suspension in this case was in the face of potentially having to release 'insanity acquittees' into the community, some of whom may well have been a danger to themselves or others.

[48] *Bedford* (n 4) 717.

[49] ibid 715.

[50] Peter Hogg, 'Necessity in a Constitutional Crisis' (1985) 15 Monash U L Rev 253; Peter Hogg and Cara Zwibel, 'The Rule of Law in the Supreme Court of Canada' (2005) 55 UTLJ 715.

sion of declarations has become increasingly common in response to findings of unconstitutionality.[51] It is also important to note that the effect of a suspension of a declaration of invalidity has been described as 'breath[ing] life' into the impugned provision for so long as the declaration remains effective.[52] This suggests that the suspended declaration acts as a source of legitimation and validity for a law that has otherwise lost that status because of a conflict with a constitutional norm. It is not clear how a creation of law (the Supreme Court) that obtains its power from the Constitution can accomplish this in a way consistent with its own constitutional limitations, but there can be no doubt that the practice is well-established even if such theoretical objections are well-founded.

4.4.2. The Proliferation of Suspended Declarations: *Bedford* and *Carter*

The use of suspended declarations of unconstitutionality raises some concerns.[53] As Niblett has observed, even in the wake of the *Schachter* criteria, the Supreme Court suspended declarations in cases that would not have been covered by those criteria.[54] He gives three such examples: *Dunmore v Ontario*, which concerned the right of agricultural workers to associate freely,[55] *Figueroa v Canada*, which concerned the restriction of benefits to parties that nominate over fifty candidates,[56]

---

[51] Bruce Ryder, 'Suspending the Charter' (2003) 21 Supreme Court Law Review 267; Grant Hoole, 'Proportionality as a Remedial Principle: A Framework for Suspended Declarations of Invalidity in Canadian Constitutional Law' (LLM thesis, 2010).

[52] *R v Moazami* 2014 BCSC 261 [22]; this analysis was alluded to, though not necessarily endorsed, in: *R v Al-Qaysi* 2016 BCSC 937 [19].

[53] In particular, Robert Leckey has been a recent vocal critic of the suspended declaration: Robert Leckey, 'Suspended Declarations of Invalidity and the Rule of Law' (*UK Const Law Blog*, 14th March 2014) ⟨https://ukconstitutionallaw.org/2014/03/12/robert-leckey-suspended-declarations-of-invalidity-and-and-the-rule-of-law/⟩ accessed 21st March 2017; Robert Leckey, 'The Harms of Remedial Discretion' (2016) 14 ICON 584; Robert Leckey, 'Remedial Practice Beyond Constitutional Text' (2016) 64 Am J Comp L 1; Robert Leckey, 'Enforcing Laws That Infringe Rights' [2016] PL 206.

[54] Anthony Niblett, 'Delaying Declarations of Constitutional Invalidity' in Frank Fagan and Saul Levmore (eds), *The Timing of Lawmaking* (Edward Elgar 2017) 305.

[55] *Dumore v Ontario* 2001 SCC 94 (CanLII), [2001] 3 SCR 1016.

[56] *Figueroa v Canada (Attorney General)* [2003] 1 SCR 912 (SCC).

and *Nguyen v Quebec*, which concerned restrictions on eligibility for minority language education.[57] None of these cases pose the extreme threat to the rule of law envisaged in the *Schachter* criteria, and yet a suspended declaration was awarded in each of them.

Leckey has drawn attention to two recent cases in which declarations of invalidity were suspended:[58] *Canada v Bedford*[59] and *Carter v Canada*.[60] The first of these cases concerned the constitutionality of criminal provisions on prostitution. It is striking how broad the rule of law analysis was in *Bedford* when compared to, say, the *Manitoba Language Rights* reference. The court considered factors such as sex workers' right to security of the person, and the concern of Canadian citizens more generally over leaving prostitution unregulated.[61] Perhaps surprisingly, it seemed to prioritise the 'great concern'[62] of Canadian citizens regarding the regulation of prostitution over the security of the person of prostitutes themselves. The Court preferred to leave *in situ* a situation that it had accepted was harmful to sex workers based on this 'great concern'. The Court suspended its declaration for a period of a year. Mouland has considered the difficulties of *Bedford* in detail.[63]

Using a disaggregated 'concern of the citizens' standard to measure the justification for suspension sets the bar almost at naught when fundamental human rights are engaged. It is difficult to envisage any human rights concern the regulation of which would not engage significant public interest, particularly once it has reached the apex of the judicial system. It is little wonder, therefore, that critics have observed a drift from the view expressed in *Schachter* that suspended declarations should be exceptional.

---

[57] *Nguyen v Quebéc (Education, Recreation and Sports)* [2009] 3 SCR 208 (SCC).

[58] Robert Leckey, 'Enforcing Laws That Infringe Rights' [2016] PL 206, 209.

[59] *Canada (Attorney General) v Bedford* 2013 SCC 72, [2013] 3 SCR 1101.

[60] *Carter v Canada (Attorney General)* 2015 SCC 5, [2015] 1 SCR 331.

[61] *Bedford* (n 59).

[62] ibid [167].

[63] Carolyn Mouland, 'Remedying the Remedy: *Bedford*'s Suspended Declaration of Invalidity' (2018) 41 Manitoba L J 281.

In *Carter*, which concerned an unconstitutional criminal prohibition of assisted suicide, the court again suspended its declaration for a year. In both this case and *Bedford*, a high degree of deference to legislative choice on policy issues is evident on the part of the Supreme Court. Leckey attributes this to a shift from an initial attitude of 'constitutional enforcement'—wherein judges will give primacy to the constitution and thus treat rights-infringing legislation as invalid—to one of 'legislative engagement', which reflects a lesser concern with the Constitutional text *per se* and a greater focus on managing a functional relationship with other aspects of government and encouraging them to pursue a constitutionally-compliant agenda.[64] Whatever one thinks of this position, it seems ill-justified by the Canadian constitutional text, and Leckey has made the point that it is not compatible with the strong affirmations of constitutional supremacy described above.[65] Other, more general, issues with suspended declarations are considered later in chapter 10.

4.5. THE REVIVAL OF UNCONSTITUTIONAL LAW

In Canada there has been little direct discussion of the revival of unconstitutional laws. There has been a confirmation of the relatively trivial point that an appellate court can reverse a finding of unconstitutionality from a lower court, which effectively 'revives' the statute.[66] It has also been held by the Supreme Court that causes of action arising from the application of Charter rights are to be dated back to the date of passing of the Charter.[67] The Supreme Court of British Columbia held in *Isaac v British Columbia (Superintendent of Motor Vehicles)* that a finding of unconstitutionality was distinct from the repeal of legislation, and in the course of so finding made the more broad statement that '[t]he concept of relying on or of reviving, in a sense, legislation that has been determined to be unconstitutional so as to fill some

[64] Leckey, 'Enforcing Laws That Infringe Rights' (n 58) 210.

[65] Robert Leckey, 'The Harms of Remedial Discretion' (2016) 14 ICON 584, 602–03.

[66] *Saltspring Island Water Preservation Society v Rockliffe* 1993 CanLII 1313 (BC CA), [1993] 4 WWR 601 [6].

[67] *Ravndahl v Saskatchewan* 2009 SCC 7 (CanLII), [2009] 1 SCR 181 [23].

gap in a legislative scheme is without basis'.[68]

That unconstitutionality does not entail repeal is illustrated by the recent complaint in Canada regarding so-called 'zombie laws', which remain on the Criminal Code even after they have been stricken by the courts. One such law was erroneously used to convict a man of second-degree murder.[69] This has led to pledges from the Minister for Justice to clear these laws from the statute book.[70] It therefore seems possible that the Canadian courts could find that unconstitutional law may be 'revived', as it remains on the statute book, but this would be a poor fit with the *dicta* in *Isaac* above and more generally with other theoretical presumptions of unconstitutionality in Canada.

4.6. CONCLUSION

In conclusion, Canada's practice of judicial review of legislation under the Charter was largely shaped by an early confrontation with a difficult scenario in the *Manitoba Language Rights* reference. This has dominated development since, as reflected in the *Schachter* case, which remains the Court's leading statement on constitutional remedies. The suspended declaration of unconstitutionality has given the Supreme Court scope to avoid significant upsets that might otherwise immediately result from a declaration of unconstitutionality.

That the Court worries about these upsets in the first place suggests an implicit assumption that unconstitutionality ineluctably means that legislation is void. This seems to be the interpretation given to the phrase *of no force or effect* in section 52 of the Charter. The practice prior to the Charter had reflected an understanding that judicial review of legislation voids the impugned legislation, and the interpretation

---

[68] *Isaac v British Columbia (Superintendent of Motor Vehicles)* 2014 BCSC 1608 (CanLII) [52].

[69] *R v Vader* 2016 ABQB 309 (CanLII).

[70] Julia Wong and Phil Heidenreich, 'Federal justice minister moves to end "zombie laws," clean up Criminal Code' (*Global News*, 8th March 2017) ⟨https://globalnews.ca/news/3295578/federal-government-expected-to-remove-zombie-laws-from-criminal-code/⟩ accessed 3rd December 2017.

of the language of the Charter has understandably been interpreted in light of this prior practice. A related concern is the Canadian judiciary's general trend of presenting unconstitutionality as a process that the Constitution autonomously regulates. Grellette observes that there are two aspects of section 52(1) of the Charter that support this.[71] The first is that the Courts have consistently found legislation void *ab initio* pursuant to review under this section. The second is that the section refers to 'any law that *is* inconsistent with the provisions of the constitution'. In other words, it does not require that the section be *applied* to a set of facts to test legal validity; rather, it simply *is* a condition of legal validity *ex proprio vigore*. As Grellette concludes, '[t]aken together, these two features of [section] 52(1) go some way towards demonstrating that the Canadian constitution is written and adjudicated upon *as though its norms are sufficient, in and of themselves, to constrain the content of valid law*'.[72]

A particularly important lesson to draw from the Canadian experience is the difficulties incurred by suspending declarations of unconstitutionality. Within the jurisdiction selection parameters of this thesis, the Supreme Court of Canada was the first apex court to suspend a declaration of unconstitutionality. Since this initial finding, the frequency of suspension has increased, and it could now justifiably be seen as the norm rather than the exception. The issues that arise as a result of suspending declarations of unconstitutionality are considered more fully in chapter 10. Canada, along with South Africa, considered in chapter 7, is an important case study jurisdiction as regards the benefits and disbenefits of this remedy and its contribution, if any, to our understanding of what courts do when they find legislation unconstitutional and the limits of that practice.

[71] Matthew Grellette, 'Legal Positivism and the Separation of Existence and Validity' (2010) 23 Ratio Juris 22, 27.

[72] ibid 27.

5 | Ireland

CHAPTER OVERVIEW

5.1. Introduction 119

5.2. Doctrines Avoiding or Curtailing Unconstitutionality 120

5.2.1. The Presumption of Constitutionality and Double Construction 120

5.2.2. Severance 121

5.3. Unconstitutionality in the Irish Courts 122

5.3.1. Early Case Law: *Murphy v Attorney General* 122

5.3.2. Drifting from the *Murphy* Standard 125

5.3.3. Completing the Drift: *A v Governor of Arbour Hill Prison* . 128

5.4. The Modern Position on Declarations of Unconstitutionality . . 131

5.5. Suspended Declarations of Unconstitutionality 134

5.6. Pre-Constitution Law 136

5.7. Revival of Unconstitutional Law 139

5.8. Conclusion 141

5.1. INTRODUCTION

The Irish Constitution is the first of two constitutions studied in this thesis that deals with the problem of 'pre-constitution law'. This is the general label I give to law that preceded the adoption of a Constitution declaring an independent and sovereign state. All the jurisdictions considered in this thesis are former colonies of the United Kingdom, and so each of them was at one time or another subject to law made by the Parliament in Westminster and/or a devolved legislature. Some countries have

transitioned more or less peacefully from this colonial status, others less so. Two of the jurisdictions studied in this thesis—Ireland and India—draw attention to the transition and applicability of pre-constitution law in their constitutional texts. This is one of the ways in which Irish practice draws attention to some peculiar points in its practice of unconstitutionality.

This distinguishing feature aside, Ireland's practice of unconstitutionality broadly mirrors the Canadian practice described in the preceding chapter, with the exception that it does not (yet) recognise the suspended declaration of unconstitutionality. Some recent cases in the Court of Appeal suggest that this may be changing but, as of time of writing, this practice has not been expressly endorsed by the Supreme Court.

5.2. DOCTRINES AVOIDING OR CURTAILING UNCONSTITUTIONALITY

5.2.1. The Presumption of Constitutionality and Double Construction

The Irish courts developed a doctrine of 'double construction' in *McDonald v Board na gCon*.[1] This is a rule whereby if there are two competing interpretations of a statute, where one is constitutional and the other is not, the court will give effect to the constitutional reading. Kenny has observed that this rule has certain limitations in Ireland.[2] Both interpretations of the legislation do not need to be *equally* plausible,[3] but the possibility of alternative readings must be 'reasonably open' to the court.[4] The ambit of the rule was curtailed in *Re Haughey*, where the court emphasised that they could not 'do violence to the plain meaning of words'.[5] Other more recent cases

[1] *McDonald v Bord na gCon* [1965] IR 217 (SC).

[2] David Kenny, 'The Separation of Powers and Remedies: The Legislative Power and Remedies for Unconstitutional Legislation in Comparative Perspective' in Eoin Carolan (ed), *The Constitution of Ireland: Perspectives and Prospects* (Bloomsbury Professional 2012).

[3] *East Donegal Co-Operative Livestock Market v Attorney General* [1970] IR 317 (SC) 341.

[4] *McDonald v Bord na gCon* (n 1) 239.

[5] *In Re Haughey* [1971] IR 217 (SC).

have also applied the rule in a more circumscribed and conservative way.[6]

5.2.2. Severance

The Irish Courts have recognised a doctrine of severance. However, it is subject to substantial limitations that restrict its utility as a limiting doctrine. Severance was first analysed by the Irish courts in *Deaton v Attorney General*.[7] The test to be applied was laid out by Fitzgerald CJ:

> [I]f what remains is so inextricably bound up with the part held invalid that remainder cannot survive independently, or if the remainder would not represent the legislative intent, the remaining part will not be severed and given constitutional validity.[8]

So, the law must be able to survive independent of the deleted words or phrase, and it must also continue to reflect the legislative intent behind the original enactment. This test has been interpreted quite restrictively by the Supreme Court.[9] A particularly stark indication of the restrictiveness of this test is *Maher v Attorney General*.[10] In this case the difficulty was with s 44(2)(a) of the Road Traffic Act 1968. This provision allowed for the admission of a certificate of blood alcohol level as *conclusive* evidence for the purpose of certain road traffic offences. The difficulty here was that the weighting of the evidence was being excluded from the ambit of the court. This could have been remedied simply by deleting the word 'conclusive'. Nevertheless, the Supreme Court refused to sever the offending word, maintaining that it would frustrate the legislative intention if they were to do so.[11]

The limitations on severance are substantial and the test laid down in *Deaton*

---

[6] *Hegarty v O'Loughran* [1990] 1 IR 148; *Grealis v Director of Public Prosecutions* [2001] 3 IR 144 (SC); *Cummins v McCartan* [2005] 3 IR 559.

[7] *Deaton v Attorney General* [1963] IR 170.

[8] ibid 147.

[9] *Blake v Attorney General* [1982] IR 117 (SC); *An Blascaod Mór Teoranta v Commissioners of Public Works (No 3)* [2000] 1 IR 6 (SC).

[10] *Maher v Attorney General* [1973] IR 140 (SC).

[11] ibid 149.

is an onerous one to meet.[12] In general, therefore, severance does not make up a significant aspect of constitutional praxis in Ireland.

5.3. UNCONSTITUTIONALITY IN THE IRISH COURTS

5.3.1. Early Case Law: *Murphy v Attorney General*

The first major case on declarations of unconstitutionality in Ireland was *Murphy v Attorney General*.[13] The plaintiffs here were a married couple. For the purposes of income tax, the wife's income was deemed to be her husband's income. The net problem was that the couple were being taxed more because they were married, and they would have enjoyed a lesser tax liability were they both single. This was held to be an unconstitutional discrimination against married couples.

In the Supreme Court, two competing views of unconstitutionality emerged, which can broadly be attributed to O'Higgins CJ and Henchy J. Henchy J rested his conception of the effects of invalidity on an *ultra vires*-type argument, echoing the submissions of counsel for the plaintiffs that Articles 15 and 24 of the Constitution required that exercising the legislative power in a manner repugnant to the Constitution was simply *ultra vires* the Oireachtas.[14] This view was taken to compel the conclusion that unconstitutional legislation must be taken to be void *ab initio*.[15] Henchy J mollified the potentially severe outcomes of this position when he conceded that 'public policy' doctrines such as *res judicata* could operate even despite a finding of voidness to prevent retrospective relief being claimed.[16]

---

[12] However, this is not to say that severance is impossible in the Irish courts. For example, in *Desmond v Glackin (No 2)* the Supreme Court deleted text from a provision of the Companies Act 1990: *Desmond v Glackin (No 2)* [1993] 3 IR 67 (SC).

[13] *Murphy v Attorney General* [1982] IR 241 (SC).

[14] ibid 290–91.

[15] Though it was not cited in *Murphy*, the Supreme Court had previously made some strong statements regarding voidness *ab initio* in a non-constitutional context in *Thomas Hunter Ltd v James Fox and Co Ltd* [1966] IR 520 (SC) 555. Voidness *ab initio* was therefore known to the Irish courts, even before Henchy J's judgment in *Murphy*.

[16] *Murphy* (n 13) 314.

O'Higgins CJ grounded his dissent on the view that since Article 25.4.1° of the Constitution prescribes that a Bill becomes law upon signing by the President (see chapter 2), then it follows that—in order to reconcile this provision with Art 15.4.2°—a declaration of invalidity may only crystallise on the date it is handed down by the relevant court, as the Constitution seems to count the Bill as law from the date of the President's signing.[17]

There are significant problems with O'Higgins CJ's dissenting view. To begin with, he makes the curious statement that '[i]rrespective of what repugnancy may exist, what has been signed, enacted and promulgated is by virtue of Article 25 immediately in force as a law of the State'.[18] This, as Henchy J notes in his first point above, does nothing at all to address the relevant *ultra vires* concerns. Fundamentally, Henchy J and O'Higgins CJ are arguing over what the necessary and sufficient conditions are for something to be considered Irish law. Henchy J's view is that the passing of a constitutionally consistent Bill by Parliament, and its subsequent assent by the President, are independently necessary and jointly sufficient conditions for a Bill to be considered valid law. O'Higgins CJ's view seems to be that the signing of a Bill by the President alone is necessary and sufficient for it to be valid Irish law.

O'Higgins CJ cannot have thought that the passing of a (constitutionally consistent) Bill in Parliament and the signing of the Bill by the President are individually necessary and jointly sufficient conditions for something to be valid Irish law. This would mean that a failure of either one condition or the other would mean the Bill was not law. That is inconsistent with his claim above that once signed by the President, the Bill is law *no matter what other flaws it bears*. It is fair to say that the reference in Art 25 to something being 'passed by both Houses of the Oireachtas' does not necessarily imply that it has been *validly* passed by the legislature, but it would be a skewed interpretation of the Constitution to bolster such a procedural provision while significantly curtailing the restriction that the legislative power be

[17] ibid 298–99.

[18] ibid 299.

exercised compatibly with other constitutional provisions (eg, fundamental rights), as per the text of Article 15. Additionally, Article 13.3.2° of the Constitution makes it clear that the President has no choice in the matter when it comes to signing a Bill;[19] if this is true, it renders the Article 25 provisions straightforwardly procedural. It would be odd to hang significant normative implications (the validity of a Bill) *solely* on a provision of this nature.

There is also an element of circularity in the Chief Justice's argument. He appears to argue that for Article 25 to be fully effective then Article 15 must be interpreted as invalidating law with prospective effect only. This argument seems to beg the question. Why think that Article 25 has any effect with respect to laws that are repugnant to the Constitution (*contra* Art 15)? O'Higgins CJ must assume his conclusion—that laws are always valid upon signing by the President—which is precisely the question in issue. He continues this question-begging when he refers to the rules being obeyed as 'laws', again the very point at issue.[20] Being less pedantic—and substituting 'law(s)' here for something like 'rule(s)' to eliminate the question-begging—O'Higgins CJ must consider that something other than legal validity is at play when a law is effective in the State. This seems particularly clear when he suggests that pre-1937 laws may be valid (in the pre-1937 Irish Free State legal system) but not *applicable* or *effective* in the Irish legal system if they fail to pass muster under Article 50.[21]

In any event, regardless of the difficulties with O'Higgins CJ's counterpoints, the *Murphy* case, through the majority judgment of Henchy J, set down a clear stance concerning the effects of a declaration of unconstitutionality in Ireland: unconstitutional legislation is to be treated as *ultra vires* the legislature and void *ab initio*. However, this stance has undergone some attrition and attenuation in subsequent cases. Quite soon after *Murphy*, the Supreme Court had to face up to the consequences of

[19] 'The President shall promulgate every law made by the Oireachtas.'.

[20] *Murphy* (n 13) 300.

[21] ibid 297.

its holding in *Muckley v Ireland*.[22]

In response to the *Murphy* decision, the Oireachtas enacted the Finance Act 1980. Section 21 of this Act provided for tax to be collected from married couples (for the years prior to 1979/80), which would not be less than their tax liability under the Income Tax Act 1967 stricken in *Murphy*. The plaintiffs in *Muckley* were a married couple assessed under this tax provision. They successfully argued that since section 21 was a reinstatement of a tax already found to be unconstitutional in *Murphy* that it, too, should be struck down.

In striking down the legislation, however, the court made some interesting remarks on the retrospectivity of the *Murphy* decision. Barrington J remarked that '[t]axes which have been paid under the impugned provisions cannot, generally, be recovered but this is not because the impugned provisions ever had any validity but because of the unfortunate fact that it is not possible to rectify all the injustices of life'.[23] This suggests that the result in *Murphy* was that void legal measures are legally ineffective but that cases involving restitution of taxation were exceptional (neither the Murphys nor the Muckleys were given restitution).

It is also worth noting that the transactions concerned in *Muckley* had all taken place *before* the *Murphy* ruling. That the plaintiffs in *Muckley* were still able to argue that their overpayments were unconstitutional on foot of *Murphy* suggests that the general rule is that declarations of unconstitutionality are to be interpreted as operating fully retroactively, but there may be exceptions to this where exceptional practical difficulty presents itself.

5.3.2. Drifting from the *Murphy* Standard

Later cases, however, do not appear to maintain the *Murphy* principle. In *McDonnell v Ireland*[24] the plaintiff sought to avail of the earlier ruling in *Cox v Ireland*,[25] where

[22] *Muckley v Ireland* [1985] IR 472 (SC).

[23] ibid 477.

[24] *McDonnell v Ireland* [1998] 1 IR 134 (SC).

[25] *Cox v Ireland* [1992] 2 IR 53 (SC).

the Supreme Court struck down section 34 of the Offences Against the State Act 1939. Section 34 provided that whenever a person holding an office in the civil service was convicted of a scheduled offence, such as membership of an unlawful organisation, such a person would immediately forfeit their office. In 1974, the plaintiff was convicted of a scheduled offence and was deemed to have forfeited his position in the civil service. Following his release from prison, he applied for reinstatement in 1975 and again in 1984. He was refused both times. In 1991, the Supreme Court declared section 34 unconstitutional in *Cox*. McDonnell then alleged that his dismissal was of no legal effect as it had been done on foot of a section found to have been void *ab initio*.

In the High Court, McDonnell's action was dismissed on the basis that the action for breach of a constitutional right was a tortious one,[26] and so section 11(2) of the Statute of Limitations 1957, which provided that an action founded on tort cannot be brought six years after the date on which the cause of action accrued, barred his action. Thus, even though the actions of the State were unconstitutional, it was now immune from suit.

The plaintiff then appealed to the Supreme Court. His appeal was rejected, and most of the members of the court held against McDonnell on the basis of *laches* or the Statute of Limitations. However, O'Flaherty J made some *obiter* remarks that suggest dissatisfaction with the *Murphy* rationale. He claimed that since the stricken section 34 of the Offences Against the State Act 1939 was in place at the time the plaintiff was prosecuted, he could not now avail of its extirpation as grounding a cause of action.[27] The minority holding of O'Higgins CJ in *Murphy* supports this, but it does not seem to square with the majority ruling. While it is true that Henchy J confined relief to the plaintiffs in the case,[28] and Griffin J made reference to the memorable phrase that 'the egg cannot be unscrambled',[29] this was still on the under-

[26] *McDonnell* (n 24) 139.

[27] ibid 142.

[28] *Murphy* (n 13) 324.

[29] ibid 331.

standing that the ineffectiveness of null legal instruments is the basic stance, which is being exceptionally deviated from in *Murphy* because of factual difficulties. This understanding also comes through in *Muckley*.

Nevertheless, O'Flaherty J went on to claim that '[t]he correct rule must be that laws should be observed until they are struck down as unconstitutional'.[30] In so doing, he relied on the Article 25 argument made by O'Higgins CJ in *Murphy*, saying that from the day the President signs a law all citizens are required to tailor their conduct according to it. O'Flaherty J made the point that every Irish judge takes an oath of office to uphold the Constitution and the laws, and judges cannot have mental reservations to apply all laws, except those that are unconstitutional.[31] There are at least two problems with this: (1) O'Flaherty J is here counting things as 'laws' that were, at least on their face, made in breach of the legislative power contained in Article 15, and (2) a judge's oath requires them to uphold both Constitution and law. Even if unconstitutional legal norms are considered 'laws' for this purpose (and this is debatable), upholding the Constitution surely requires more attention to the language of Article 15, which is not mentioned even once in O'Flaherty J's judgment. As with O'Higgins CJ before, there is not much time spent on the *ultra vires* concerns that motivated Henchy J in his judgment in *Murphy*. Despite these difficulties, O'Flaherty J concluded that '[t]he consequence of striking down legislation can only crystallise in respect of the immediate litigation which gave rise to the declaration of invalidity'.[32] Crucially, this would become a running theme in the later *A* case, despite the tension between this statement and what *Murphy* originally held.

In any case, the drift from *Murphy* was only nascent here. The outcome of *McDonnell* could still explain and achieve the same result as in *Murphy*, albeit with a tonal shift towards the minority position in that case. However, as Doyle has noted, the reasoning in *McDonnell* could not support the result in *Muckley*.[33] O'Flaherty

[30] *McDonnell* (n 24) 143.

[31] ibid 143.

[32] ibid 144.

[33] Oran Doyle, *Constitutional Law: Text, Cases and Materials* (Clarus Press 2008) 451.

J's reasoning would block the Muckleys from asserting their claim as the declaration of unconstitutionality did not crystallise in their claim; it crystallised in *Murphy*.

5.3.3. Completing the Drift: *A v Governor of Arbour Hill Prison*

A major reconsideration of the foundational principles of the Irish stance on declarations of unconstitutionality[34] arose in the case of *A v Governor of Arbour Hill Prison*,[35] which has been cited in the UK Supreme Court,[36] the Scottish Court of Session (Inner House)[37] and the Supreme Court of Papua New Guinea.[38] This case arose from an earlier decision of the Supreme Court in *CC v Ireland*.[39] In that earlier decision, the Court declared section 1(1) of the Criminal Law (Amendment) Act 1935—providing for the offence of unlawful carnal knowledge—to be inconsistent with the Constitution, as it did not provide for a defence of reasonable mistake as to the age of the complainant. This was the section under which A was imprisoned. Thus, A instituted *habeas corpus* proceedings challenging his incarceration.

The High Court initially ordered the release of A. On appeal the High Court decision was overturned by the Supreme Court *ex tempore*.[40] Murray CJ raised the point that because 'no interpretation of the Constitution is intended to be final for all time'.[41] Having earlier made the point that judges inevitably change the law,[42] he

---

[34] Though sadly not of the inconsistencies between *Muckley* and *McDonnell* discussed above.

[35] *A v Governor of Arbour Hill Prison* [2006] IESC 45, [2006] 4 IR 88.

[36] *Cadder v Her Majesty's Advocate* [2010] UKSC 43, [2010] 1 WLR 2601.

[37] *Graham Gordon (Motion for Judicial Review)* [2013] ScotCS CSIH 101 [3].

[38] *Polem Enterprise Ltd v Attorney General* [2008] PGSC 9, (2009) 1 LRC 627, 46–50.

[39] *CC v Ireland* [2006] IESC 33, [2006] 4 IR 1.

[40] Cf the result arrived at by the House of Lords in *R v Governor of Brockhill Prison, ex parte Evans* [2000] UKHL 48, [2001] 1 AC 19. The cases are different in that Evans was looking for damages for wrongful imprisonment and not *habeas corpus* as Evans' time had already been fully served. Nevertheless, it is interesting to note that the House of Lords had little sympathy for the good faith reliance on invalid acts invoked to detain Evans. Cf also *Percy v Hall* [1997] QB 924 (HC), where there was no apparent problem with arrests being made in good faith on invalid by-laws

[41] Here citing Walsh J in *McGee v Attorney General* [1974] IR 284 (SC) 319.

[42] Here citing himself: 'first, there is the law; then there is interpretation. Then interpretation is the law.' *Crilly v T & J Farrington Ltd* [2001] 3 IR 251 (SC) 286.

suggested that it may be the case that a statute that passed muster in the 1940s or 50s could be found invalid by reference to 21st century shifts in value.[43] After considering comparative approaches in Canada, the United States, and the European Union, he rejected the notion that the law admits of 'absolute retrospectivity' and so took it that the prior cases on Ireland and the Irish Constitution could not be taken to impute this proposition. Of course, mere quantitative weight in favour of a theoretical position is not particularly robust evidence for this conclusion, and in the field of constitutional law ethnographic differences can be decisive.[44]

Denham J (McGuinness J concurring) took the somewhat surprising view that the general principle in Ireland was that a declaration of invalidity of a law applies to the parties of the relevant litigation retrospectively, and prospectively only for the rest of society. If this were true, then *Muckley* could never have occurred. It is also questionable the extent to which this reflects the view of Henchy J in the *Murphy* case, which the Court purported to follow. Henchy J originally claimed that '[o]nce it has been judicially established that a statutory provision enacted by the Oireachtas is repugnant to the Constitution, and that it therefore incurred invalidity from the date of its enactment, the condemned provision will normally provide no legal justification for any acts done or left undone, or for transactions undertaken in pursuance of it'.[45] This seems to run contrary to Denham J's view.

Hardiman J usefully distilled the prior Irish case law down to the following five propositions:[46]

> (1) Post-Constitution statutes that are found invalid are to be held invalid from the date of their enactment.
> (2) Not all legal action taken on foot of invalid statutes is null.[47]
> (3) Considerations such as economic necessity, practical convenience, public policy, and the equity of the case may all require effect to be given to certain action taken

[43] *A v Governor of Arbour Hill Prison* (n 35) [84].

[44] I return to the issue ethnographic differences in unconstitutionality in chapter 11.

[45] *Murphy* (n 13) 313.

[46] *A v Governor of Arbour Hill Prison* (n 35) [236].

[47] *The State (Byrne) v Frawley* [1978] IR 326 (SC) 314–15; *Murphy* (n 13) 307; *McDonnell* (n 24) 143.

> under a void statute.[48]
> (4) This effect is can be justified: (a) by the difficulty of undoing the past; (b) the compulsion of public order and the common good; (c) acquiescence of the right-holder.[49]
> (5) The court should not attempt to lay down a rigid rule as to what the effect of an unconstitutional statute might be.[50]

Hardiman J particularly seized upon (2) and distinguished between the voidness of a statute and the voidness of action taken under it.[51] Nowhere does Hardiman J turn his mind to the much more knotted question of what the legal basis for such action might be. He rhetorically amplifies aspects of Henchy J's judgment in *Murphy* that would support this conclusion while also downplaying those parts that would suggest that there was in fact no law to continue to justify A's detention.

All five judges of the court rely on similar cases in support of their views, and in the process miss several that point to difficulties or inconsistencies. Doyle has drawn attention to the *Muckley* problem already discussed in the context of Denham J's judgment.[52] Additionally, Ring[53] has drawn a contrast between the handling of the application in *A* and the judgment of the Supreme Court in *Glavin v Governor of Mountjoy Prison*.[54] The situation in *Glavin* arose from the judgment of the Supreme Court in *Shelly v Mahon*,[55] wherein the respondent was a District Court judge who was past the retirement age at the time he tried and convicted the applicant. This made the applicant's conviction unconstitutional, and so their conviction was quashed. Following the decision in *Shelly*, the applicant in *Glavin*—who had been sent forward for trial in the Circuit Court by the same judge—also applied for an order of *certiorari*. This order was also granted. Although certain particulars

---

[48] *Murphy* (n 13) 321; *Chicot County Drainage District v Baxter State Bank* 308 US 371, 374 (1940).

[49] *Murphy* (n 13) 307, 314–15, 321.

[50] ibid 315.

[51] *A v Governor of Arbour Hill Prison* (n 35) [237].

[52] Doyle, *Constitutional Law: Text, Cases and Materials* (n 33) 451.

[53] Sinéad Ring, 'Collateral Challenges to a Declaration of Unconstitutionality: *A v Governor of Arbour Hill Prison*' (2006) 14 ISLR 245, 253.

[54] *Glavin v Governor of Mountjoy Prison* [1991] 2 IR 421 (SC).

[55] *Shelly v Mahon* [1990] IR 36 (SC).

of the cases differ—the seriousness of the crimes at issue, the fact that *A* was a collateral attack and *Glavin* was not—the results of the cases diverge completely.[56] The magnitude of this divergence is not particularly proportionate to the factual differences between the cases.

The foregoing analysis suggests that the Irish case law on the underlying theory of this area is quite confused. Inconsistency emerges between the rationale initially articulated in *Murphy* and reaffirmed in *Muckley* and the subsequent tacit shift in *McDonnell*, which was reinforced by the Supreme Court in *A*. This shift in rationale seems to suggest that while the original position identified in *Murphy* continues to be cited to, actual practice no longer reflects the theory established in that case.

5.4. THE MODERN POSITION ON DECLARATIONS OF UNCONSTITUTIONALITY

The post-*A* case law has not revisited in detail the theoretical commitments underpinning declarations of unconstitutionality, and it can be dealt with more briefly. In *Damache v DPP*[57] the Supreme Court struck down section 29(1) of the Offences Against the State Act 1939, a provision pertaining to search warrants. This had the effect of potentially disrupting both secured convictions and prosecutions in train. This threat of disruption lead to several further cases being taken in the Court of Criminal Appeal. In *People (DPP) v Cunningham*[58] the Court allowed the appeal of the applicant on the basis that where there was still a right of appeal then finality did not attach to the decision and so it was not *res judicata*. In *People (DPP) v Hughes*[59] the Court refused an application because the accused had pleaded guilty and leave to appeal was sought out of time. In general, the post-*Damache* case law established

[56] Fanning has drawn attention to a similar tension between *Grealis v Director of Public Prosecutions* [2001] 3 IR 144 (SC), where the Supreme Court found that the Non-Fatal Offences Against the Person Act 1997 abolished the common law crimes of assault and battery, and *A*. In *Grealis* there seemed to be no concern about windfall benefits falling to defendants in ongoing prosecutions: Rossa Fanning, 'Hard Case; Bad Law? The Supreme Court Decision in *A v Governor of Arbour Hill Prison*' (2005) 40 Ir Jur 188, 209.

[57] *Damache v Director of Public Prosecutions* [2012] IESC 11, [2012] 2 IR 266.

[58] *The People (DPP) v Cunningham* [2012] IECCA 64, [2013] 2 IR 631.

[59] *The People (DPP) v Hughes* [2012] IECCA 69, [2013] 2 IR 619.

three conditions for an accused to be able to rely on the unconstitutionality of section 29(1): (1) the accused must have raised the issue themselves at trial, and (2) the accused must not have taken steps that could be taken as acquiescing to the charge, such as pleading guilty, and (3) the proceedings must be non-final; there must be a further trial or appeal pending.[60] The Supreme Court approved this test in *Director of Public Prosecutions v Doyle*.[61]

The requirement that the accused themselves must have raised the issue at trial to benefit from a declaration of unconstitutionality handed down in separate proceedings confirms that 'windfall' benefits will be difficult, if not impossible, for an accused to secure where they are completely unaware of any constitutional infirmity and have not tried to argue one.[62] Again, this is difficult to square with the nominally *erga omnes*, sweeping effects of declarations of unconstitutionality under *Murphy*. The post-*Damache* case law frequently cites to a passage from the judgment of Murray CJ in *A*:

> In a criminal prosecution where the State relies in good faith on a statute in force at the time and the accused does not seek to impugn the bringing or conduct of the prosecution, on any grounds that may in law be open to him or her, including the constitutionality of the statute, before the case reaches finality, on appeal or otherwise, then the final decision in the case must be deemed to be and to remain lawful notwithstanding any subsequent ruling that the statute, or a provision of it, is unconstitutional. That is the general principle.[63]

This 'general principle' is not the one articulated in *Murphy*. Although acquiescence was mentioned even in the earliest case law on declarations of unconstitution-

---

[60] *The People (DPP) v Kavanagh* [2012] IECCA 65; *The People (DPP) v O'Brien* [2012] IECCA 68.

[61] *The People (DPP) v Doyle* [2017] IESC 1.

[62] It is notable that this sort of difficulty ('fishing one case from the stream of appellate review, using it as a vehicle for pronouncing new constitutional standards, and then permitting a stream of similar cases subsequently to flow by unaffected by that rule') was something that Justice Harlan had objected to in his dissents against *Linkletter*, considered earlier in chapter 3: *Mackey v United States* 401 US 667, 679 (1971).

[63] *A v Governor of Arbour Hill Prison* (n 35) [125.] Cited in: *Kennedy v Governor of Mountjoy Prison* [2017] IEHC 402 [23]; *Clarke v Governor of Mountjoy Prison* [2016] IEHC 278 [37]; *Kovacs v Governor of Mountjoy Women's Prison (Dochas Centre)* [2015] IEHC 418 [30]; *Willis v Governor of Wheatfield Prison* [2015] IEHC 251; *The People (DPP) v McKevitt* [2014] IECCA 19; *The People (DPP) v Timmons* [2013] IECCA 86 [19]; *The People (DPP) v Kavanagh* [2012] IECCA 65 [54].

ality, the emphasis that has been placed on this quotation is far removed from the original rule. The effect of this principle is to eliminate unconstitutionality passively happening to benefit an accused; if an accused is to advance unconstitutionality as an element of their case, they must do themselves. At least in respect of criminal law matters, the rule is now that unconstitutionality only crystallises on a case-by-case basis and there is no presumption of any automatic retrospective effect.

In some respects, this is not completely novel. As pointed out earlier, relief was restricted to the applicants even in the early cases of *Murphy* and *Muckley*.[64] Limitations on the fallout of unconstitutionality are nothing new to the Irish courts. The pillars on which these holdings have been established have, however, shown themselves to be made of sand; moreover, this is a subtly shifting sand that has generated an insubstantial theory of the consequences of unconstitutionality following a court ruling.

In summary, inconsistency emerges between the rationale initially articulated in *Murphy* and reaffirmed in *Muckley*, and the subsequent tacit shift in understanding evidenced in *McDonnell* that was then taken up by the Supreme Court in *A* and completed in *Damache* and its progeny. While the *Murphy* case continues to be cited with a degree of nominal reverence, practice has diverged from the strong statements of a presumption of retrospective effect advanced in that case.[65]

---

[64] See also the holding of the Supreme Court in *In Re Thomondgate No 1 Area CPO* 1986 WJSC-SC 1512, refusing to extend a windfall benefit arising from *Blake v Attorney General* [1982] IR 117 (SC). However, it should be noted that this case may have rested more on the construction of section 2(2) of the Acquisition of Land (Assessment of Compensation Act) 1919 than general constitutional principle.

[65] A similar conclusion to that reached in *A* had already been achieved in the context of review of administrative action: in that context there is a presumption of validity which entails a kind of presumption of prospective effect: *Campus Oil v Minister for Industry and Energy (No 2)* [1983] IR 88 (SC) 107; *The State (Abenglen Properties Ltd) v Dublin Corporation* [1984] IR 381 (SC)
As in *A*, the Supreme Court remarked in the *O'Keeffe* case that while '[i]t is usual to say that an *ultra vires* decision is void and a nullity … it is clear that it is wrong to conclude that such decisions are completely devoid of legal consequences'. *O'Keeffe v An Bord Pleanála* [1993] 1 IR 39 (SC) 49
The most minimal interpretation of this statement is surely correct: a legally invalid decision must at least have legal consequences in grounding a judicial review action. However, in the review of the constitutionality of legislation this goes beyond mere equitable limitation doctrines; it seems more like a total abandonment of what voidness *ab initio* really means by treating retrospective alterations to the juridical status of past action as the exception rather than the norm.

5.5. SUSPENDED DECLARATIONS OF UNCONSTITUTIONALITY

It might be that the issue to be resolved here could be reconsidered as a solely remedial one: whatever the underlying theory regarding declarations of unconstitutionality, what remedies should be available when legislation is struck down? This is a fundamentally pragmatic question that, as a framing device, sidesteps the need to provide any thoroughgoing account of unconstitutionality. An element of this pragmatism can be detected in an *obiter* statement by O'Higgins CJ *Blake v Attorney General*, where the Chief Justice suggested that the courts should either adjourn or stay applications for possession following a lacuna left by the striking down of parts of the Rent Restrictions Act 1960 'in the reasonable expectation of new legislation'.[66] This would, in effect, be tantamount to the suspended declaration remedy that is employed in Canada[67] and South Africa.[68]

For some time, this suggestion was left unexpanded upon by the courts. However, in *Kinsella v Governor of Mountjoy Prison*[69] Hogan J seemed to indicate a potential shift towards the grant of suspended declarations.[70] This shift has been met with some warm academic reception in Ireland.[71] Initially there was reticence to follow up on the approach in *Kinsella*. In *Bederev v Ireland*[72] Hogan J was presented with what seemed like a prime opportunity to further develop the approach he had first articulated in *Kinsella*. In *Bederev* the issue at stake was a serious one: the Court of Appeal was asked whether section 2(2) of the Misuse of Drugs Act 1977

---

[66] *Blake v Attorney General* (n 64) 141–42.

[67] *Schachter v Canada* [1992] 2 SCR 679 (SCC).

[68] Constitution of South Africa 1996, s 172.

[69] *Kinsella v Governor of Mountjoy Prison* [2011] IEHC 235, [2012] 1 IR 467.

[70] 'The proposed solution – i.e., upholding the claim of a violation of a constitutional right, but giving the authorities an opportunity to remedy this breach – is also perhaps the one which is the most apt having regard to the principles of the separation of powers'. ibid [15].

[71] Eoin Carolan, 'The relationship between judicial remedies and the separation of powers: collaborative constitutionalism and the suspended declaration of invalidity' (2011) 46 Ir Jur 180; Kenny, 'The Separation of Powers and Remedies: The Legislative Power and Remedies for Unconstitutional Legislation in Comparative Perspective' (n 2).

[72] *Bederev v Ireland* [2015] IECA 38.

was constitutional. This section provided a power for the Government to declare certain substances 'controlled drugs' for the purposes of the 1977 Act. In that case, Hogan J again delivered judgment for the Court of Appeal but did not suspend the declaration in this case. This was ultimately rendered moot as the Supreme Court reversed the decision.

Some stronger evidence of suspension as an emerging phenomenon is now observable following the Supreme Court's decision in *NHV v Minister for Justice*.[73] This case concerned the constitutionality of the direct provision system for refugees in Ireland. O'Donnell J, giving judgment for the Supreme Court, suspended the Court's order for a period of six months.[74] The Court also seems to have endorsed this approach in *PC v Minister for Social Protection*.[75] The Supreme Court expanded briefly on this remedy in late 2017, but did not provide much in the way of clarification as to the circumstances in which it might be granted in the future. It merely noted that the case before the court in *NHV* was a very exceptional one.[76]

More recently, the Court of Appeal has issued further suspended declarations of unconstitutionality in *AB v Clinical Director of St Loman's Hospital*[77] and *Osinuga and Agha v Minister for Social Protection*.[78] The *AB* case concerned the constitutionality of a law providing for the detention of certain patients in long term psychiatric units. The *Osinuga and Agha* cases concerned the constitutionality of a law preventing the payment of child benefit to children based on of the immigration status of their parents. The consequence of a declaration of unconstitutionality in *AB* would have been that there would have been no basis for the continued detention of severely vulnerable patients that could be a risk to themselves or others.[79]

---

[73] *NHV v Minister for Justice* [2017] IESC 35.

[74] ibid [21].

[75] *PC v Minister for Social Protection* [2017] IESC 63.

[76] *NHV v Minister for Justice* [2017] IESC 82.

[77] *AB v The Clinical Director of St Loman's Hospital* [2018] IECA 123 [109]–[113].

[78] *Osinuga and Agha v Minister for Social Protection* [2018] IECA 155.

[79] In some respects this is similar to the Canadian case *R v Swain* 1991 CanLII 104 (SCC), [1991] 1 SCR 933, where the Supreme Court of Canada suspended the unconstitutionality of a law provid-

This was an alarming prospect that the Court of Appeal mollified by suspending its declaration. The consequence in *Osinuga and Agha* was less pronounced, as noted by the Court itself.[80] Here the Court was satisfied to suspend the declaration merely on the basis that straightforward invalidation would have budgetary implications for government[81] and that it could lead to an overbroad class of further beneficiaries claiming child benefit who might not have been successful themselves had they taken their own challenges.[82] Finally, the court suspended the order but specifically exempted the claimants from this suspension.[83]

These early suggestions are not yet enough to suggest that Irish law has definitely moved on from the *A*/*Damache* standard. The law, thus, remains that unconstitutionality crystallises when judgment is handed down by a court, and it crystallises immediately.

5.6. PRE-CONSTITUTION LAW

Due to Ireland's history as a colony of the UK, there are two different procedures following which a statute may be stricken for unconstitutionality. Pre-1937 statutes are negated under Article 50, which is a transitional provision providing for the carrying over of laws enacted by the Irish Free State or Westminster Parliament prior to 1937. Post-1937 statutes are voided by operation of Article 15.4 in the manner described above.

Article 50 reads as follows:

> 1. Subject to this Constitution and to the extent to which they are not inconsistent therewith, the laws in force in Saorstát Éireann immediately prior to the date of the coming into operation of this Constitution shall continue to be of full force and effect until the same or any of them shall have been repealed or amended by enactment of the Oireachtas.

ing for the detention of individuals found not guilty of a criminal charge by reason of insanity.

[80] *Osinuga and Agha v Minister for Social Protection* (n 78) [61].

[81] ibid [62].

[82] ibid [63].

[83] ibid [64].

> 2. Laws enacted before, but expressed to come into force after, the coming into operation of this Constitution, shall, unless otherwise enacted by the Oireachtas, come into force in accordance with the terms thereof.

If a provision is found 'inconsistent' with the 1937 Constitution, then it is deemed never to have carried over and thus never had effect in the Republic of Ireland. The language of 'inconsistency' here—rather than 'invalidity' or 'repugnancy'—is employed because the laws in question are assumed to be valid by reference to the power of the Parliament that enacted them. The question, rather, is whether they continue to be *applicable* in the Republic of Ireland.[84] It is interesting to note that the Irish Courts accept a view not unlike hard positivism here: the validity of a law is to be determined by reference to its sources, not its merits.[85]

There is a further complication that can occur here: do pre-1937 laws that were amended by the Oireachtas sometime after 1937 enjoy a presumption of constitutionality? If this were the case, then a validity analysis, rather than an applicability analysis, might be more appropriate. In *O'Brien v Manufacturing Engineering Co Ltd* the High Court took the view that the Oireachtas should be assumed to think that a pre-1937 law is constitutional if it makes an amendment to it.[86] In the subsequent case of *People (Attorney General) v Conmey*[87] the Supreme Court concluded that wholesale re-enactment of a pre-1937 provision by the Oireachtas would be sufficient to attract a presumption of constitutionality as well.[88]

A further development on this score arose in *An Blascaod Mór Teoranta v Commissioners of Public Works*.[89] Budd J in the High Court observed that among the legislation that was challenged by the applicant in that case, some pre-1937 Acts had

---

[84] *The State (Sheerin) v Kennedy* [1966] IR 379 (SC); *The People (DPP) v MS* [2003] 1 IR 606 (SC); *JP v Director of Public Prosecutions* [2009] 3 IR 215 (HC).

[85] *Geoghegan v Institute of Chartered Accountants in Ireland* [1995] 3 IR 86 (HC) 95 (Murphy J).

[86] *O'Brien v Manufacturing Engineering Co Ltd* [1973] IR 334 (HC) 340. The Supreme Court, on appeal, did not consider this issue expressly but Walsh J did seem to agree with the High Court: 361.

[87] *People (Attorney General) v Conmey* [1975] IR 341 (SC).

[88] A sentiment repeated in *ESB v Gormley* [1985] IR 129 (SC) 147.

[89] *An Blascaod Mór Teoranta v Commissioners of Public Works (No 3)* [1998] IEHC 38.

been incorporated into numerous items of other legislation (both pre- and post-1937).[90] On this basis, Budd J thought that the presumption of constitutionality should apply to the earlier pieces of legislation as well.

However, this view of extending the presumption of constitutionality by incorporation and reference has not been warmly received in all the Irish case law. Barrington J drew attention to some potential difficulties in this approach in *Brennan v Attorney General*:

> The issue is whether the statute, bearing the meaning it then had, is or is not consistent with the Constitution of 1937 and was or was not carried forward by Article 50 … If the statute, as so interpreted, was not carried forward by Article 50 it is immaterial that the Oireachtas later tried to amend or adapt it because there was nothing for them to amend or adapt.[91]

The idea that the legislature cannot 'amend' a nullity had received some faint attention in prior cases. Casey gives the example of the Road Traffic (Amendment) Act 1973.[92] Section 7 of this Act amended section 44(2) of the Road Traffic Act 1968 Act, but this latter section had been found unconstitutional in *Maher v Attorney General*.[93] The question that then arises is: was there anything for section 7 of the Road Traffic (Amendment) Act 1973 to amend? Given that *Murphy* means that the relevant section of the 1968 Act never had legal effect[94] it appears there was no *legal* entity for the 1973 Act to amend. Given that the section is reproduced in a Table in the amending Act here, however, this difficulty may be avoided. The reproduction in the Table might be taken as a tacit re-enactment of section 44 of the 1968 Act, which would then be amended by section 7 of the 1973 Act. This issue did not arise in any case, and the Road Traffic (Amendment) Act 1973 was repealed by the Road Traffic (Amendment) Act 1978, so it is not clear what an Irish court would have held if it had been confronted with this issue.

---

[90] *An Blascaod Mór Teoranta v Commissioners of Public Works (No 3)* (n 89) [244]–[45].

[91] *Brennan v Attorney General* [1983] ILRM 449 (HC) 479.

[92] James Casey, *Constitutional Law in Ireland* (3rd edn, Sweet & Maxwell 2000) 372–73.

[93] *Maher v Attorney General* (n 10).

[94] *In Re Article 26 and the Housing (Private Rented Dwellings) Bill 1981* [1983] IR 181 (SC) 186.

The better view is probably that incorporation by reference is theoretically tenable and sensible. Invalid norms do 'exist' in some sense. The hard-line stance on existence being coextensive with validity is one I shall attempt to cast some doubt on in this thesis, which idea I explore further in chapters 8 and 9. Without this premise to rely on, Barrington J's argument collapses.

5.7. REVIVAL OF UNCONSTITUTIONAL LAW

Hogan and Whyte have suggested that the revival of unconstitutional law must be treated with some caution outside of the United States.[95] The argument is that this practice depends heavily on a particular characterisation of the judicial power to review legislation. Recall that judicial review of legislation in the United States is a qualified power derived from *Marbury*, and that the *erga omnes* effect of declarations of unconstitutionality derives from the doctrine of precedent. This is so because there is no power under the United States Constitution to *invalidate* a law; rather, laws that are struck down by the courts are kept on the statute book, but they are not enforced because of the finding of unconstitutionality.[96] All of these are contingent features of the United States legal system, and *a fortiori* are ones that are not replicated in other legal systems. Most modern constitutional arrangements provide for an express power of *invalidation* of legal instruments by the courts.

McMahon has attempted to provide an analysis based on Irish law that would avoid the need for re-enactment of unconstitutional law.[97] He argues that to be law an Act must satisfy two conditions. It must: (1) pass through the legislative process, and (2) not be contrary to the Constitution. Where an Act is held invalid, then it is (2) that is not satisfied, but if the decision in which it was found invalid is over-

[95] Gerard Hogan and Gerry Whyte, *J M Kelly: The Irish Constitution* (4th edn, Tottel Publishing 2003) 896 fn 646.

[96] Gerard Hogan, 'Declarations of Incompatibility, Inapplicability and Invalidity: Rights, Remedies and the Aftermath' in Kieran Bradley, Noel Travers and Anthony Whelan (eds), *Of Courts and Constitutions:* Liber Amicorum *in Honour of Niall Fennelly* (Hart Publishing 2014) [16].

[97] Bryan McMahon, 'Review of JM Kelly: The Irish Constitution' (1981) 16 Ir Jur 190, 200.

ruled then it *is* satisfied. This is so only if the invalidity is taken as flowing from the earlier decision; that is, if the earlier decision was *constitutive* of the invalidity, then a later overruling decision can 'un-constitute' the invalidity, so to speak. Casey makes the point that this would gel with the observable phenomenon that a judicial finding of unconstitutionality does not repeal a statutory provision. However, a more significant barrier to this argument in Ireland is that *Murphy* and even *A* seem to assume that judicial pronouncements are *declarative* of unconstitutionality, rather than constitutive of it.

A particularly odd case in Ireland is *WM*.[98] Hogan and Whyte refer to this as a 'very special case'.[99] In this case, section 5 of the Punishment of Incest Act 1908 was held not to have been carried over by the 1937 Constitution (it provided for a mandatory *in camera* rule in incest cases, which was held to conflict with Art 34). However, its effect was restored by section 45(3) of the Courts (Supplemental Provisions) Act 1961. In some respects this is due to section 45(3) itself being something of an oddity. It provided that:

> Any provision contained in any statute of the Parliament of the former United Kingdom or of the Oireachtas of Saorstát Éireann which provided for the administration of justice otherwise than in public and which is not in force solely by reason of its being inconsistent with the provisions of the Constitution of Saorstát Éireann or the Constitution … shall have full force and effect.

This provision thus seemed to suggest that any pre-constitution law not in force that provided for an *in camera* proceeding was given new life, regardless of the reason for its not being carried over by Article 50. Thus, in *WM*, Carney J held both that the 1908 Act had not survived the passing of the 1937 Constitution and that it had been brought back into force by the 1961 Act. This seems not unlike the reading of the United States cases of *Adkins* and *Parrish* cases given by Attorney General Cummings discussed above. The two situations are not, of course, the same. *Adkins*/*Parrish* involved a situation where a judicial decision had invalidated a law,

[98] *The People (DPP) v WM* [1995] 1 IR 226 (HC).

[99] Hogan and Whyte (n 95) 849 fn 437.

which decision was itself then invalidated. This created a kind of 'double negative' that cancelled out the ineffectiveness of the law. In *WM* the statute had never been effective post-1937 at all, regardless of judicial decisions, but was revived instead by a legislative instrument.

The unusual *WM* case and McMahon's arguments aside, the idea of the revival of dead law has not arisen directly in other Irish cases. Both data run contrary to the *Murphy*/*A* orthodoxy that judges do not constitute unconstitutionality in Ireland. It therefore seems unlikely that undead laws would be anything other than highly aberrant in this jurisdiction.

5.8. CONCLUSION

In general, Ireland has adhered relatively closely to its Constitutional text as outlined in chapter 2. The reference therein to invalidity grounded an early and strong adherence to voidness *ab initio* in *Murphy*. However, gradually and subtly through the case law this rule transitioned into one where the retrospective extension of unconstitutionality became the exception rather than the norm. A significant cut-off point in this regard was *A v Governor of Arbour Hill Prison*. In that case the Supreme Court made several statements to the effect of avoiding draconian or administratively chaotic effects following unconstitutionality, and it restricted the scope of its earlier finding in the *CC* case. These statements of the Court have been taken up in later cases, further supporting the view that there is an increasing antimony in the Irish courts' practice between finding legislation *invalid* (and thus void *ab initio*) and localising the effects of unconstitutionality to a few parties.

The most recent trend in the Irish law is the embrace of the suspended declaration of unconstitutionality, beginning in *NHV* and following through to recent Court of Appeal cases. It remains to be seen how this development will take shape in Ireland, but it seems likely that it will become an increasing feature of practice.

The treatment of pre-constitution law is notable for how the Irish judiciary seem to refuse to make validity claims on such laws. Although many of these were made

obsolete by the Statute Law Revision Act 2007, a number remain in force including provisions on criminal penalties for perjury,[100] repleiving unlawfully retained property,[101] and wards of court.[102] As is evident from this list, these statutes tend to cover unusual and rarely-invoked elements of law. They could still be subject to constitutional challenge, though how successful such a challenge is likely to be after the passing of the 2007 Act is difficult to determine, as that Act could be taken as giving implicit imprimatur to the retained pieces of legislation.[103] As shall be seen in the next chapter, India has developed a similar disposition to pre-constitution law.

[100] Perjury Act 1586.

[101] Distress Act 1796.

[102] Lunacy Regulation Act 1871

[103] See the discussion of the *WM* case and the Road Traffic (Amendment) Act 1973 above.

6 | India

CHAPTER OVERVIEW

6.1. Introduction 143

6.2. Doctrines Avoiding or Curtailing Unconstitutionality 144

6.2.1. Presumption of Constitutionality and Double Construction 144

6.2.2. Severance 144

6.3. Degrees of Voidness 145

6.3.1. Voidness under Article 13(1) 147

6.3.2. Voidness under Article 13(2) 153

6.3.3. Some Clarity: *Jagannath* and *Ambica Mills* 154

6.4. The Doctrine of Eclipse 157

6.4.1. Extension of the Doctrine to Post-Constitution Law? 160

6.4.2. Restrictions on the Doctrine of Eclipse 161

6.5. Conclusion 163

6.1. INTRODUCTION

As described in chapter 2, the Indian Constitution provides what might be, in the context of this thesis, the most unambiguous textual support for an unconstitutionality doctrine. It expressly requires that unconstitutional legislation be 'void'. As shall be explored below, this idea is distinctly more relative and malleable than it may first appear. This is particularly clear in the case of the 'doctrine of eclipse', which maintains that unconstitutional law is merely overshadowed or obstructed, and that these obstructions can be removed to bring the law back into force. This doctrine is also interesting for reflecting a similar distinction in treatment between

pre- and post-constitution law, as was discussed in the context of Ireland in chapter 5.

The Indian experience serves to exemplify two important and related lessons in the study of unconstitutionality: (1) the constitutional text may not constrain practice as much as might be expected, and (2) the use of foreign precedent can generate confusion where that precedent is not adapted to the context of the domestic jurisdiction.

6.2. DOCTRINES AVOIDING OR CURTAILING UNCONSTITUTIONALITY

6.2.1. Presumption of Constitutionality and Double Construction

As in the other jurisdictions surveyed in this thesis, the Indian courts have developed multiple doctrinal responses that hem in the effect of its constitutional voidness provision. For example, there is a general presumption that legislation is constitutional.[1] However, it is not clear whether this presumption extends to pre-constitution legislation.[2] These rules are accompanied by a double construction rule whereby if multiple interpretations of a statute are possible then the statute must be read as constitutional.[3]

6.2.2. Severance

The Indian courts practice severance where invalid and valid parts of unconstitutional legislation can be separated. The leading case on this remedy is *Chamarbaugwala v Union of India*.[4] Severance may apply to both pre- and post-Constitution law.[5] In the *Chamarbaugwala* case, the Court reviewed Indian and comparative law,

---

[1] *State of Kerala v NM Thomas* [1975] INSC 224, [1976] SCR 1 906.

[2] *State of West Bengal vs Anwar All Sarkarhabib* [1952] SCR 284 (INSC) 22.

[3] *Sunil Batra v Delhi Administration* [1978] INSC 148, [1979] SCR 1 392; *Hariharan v Reserve Bank of India* [1999] INSC 37.

[4] *R M D Chamarbaugwala v Union of India* [1957] INSC 32, [1957] SCR 930.

[5] *State of Bombay v FN Balsara* [1951] INSC 38, [1951] SCR 682.

and summarised the position as follows:

> When a statute is in part void, it will be enforced as regards the rest, if that is severable from what is invalid. It is immaterial for the purpose of this rule whether the invalidity of the statute arises by reason of its subject-matter being outside the competence of the legislature or by reason of its provisions contravening constitutional prohibitions.[6]

As shall be seen below, in some other aspects of Indian constitutional law there is a distinction (that is not always clear-cut) drawn between breaches of legislative competence and breaches of certain fundamental rights.

6.3. DEGREES OF VOIDNESS

Although India has the strongest textual foundation for a practice of invalidity or voidness, the interpretation of the word 'void' in Article 13 has not reflected the understanding of that term endorsed by the Irish and Canadian courts.[7] While both the Irish and Canadian courts have found that their provisions require a default stance of voidness *ab initio*, the Indian courts have not taken this course. The voidness in the Indian courts is often not of an absolute character. It is striking that this is the case notwithstanding that the Indian Constitution is the only one that incorporates the word 'void' in its Constitutional text.[8] This voidness applies to both pre-constitution law (Article 13(1)) and post-constitution law (Article 13(2)).

In some cases, the standard applied suggests that, rather than the Constitution operating to invalidate laws, the judiciary suspends or blocks their enforcement. Consider the following statement by Das J in *Keshavan Madhava Menon v State of Bombay*,[9] one of the earliest cases to consider Article 13:

---

[6] *R M D Chamarbaugwala v Union of India* (n 4) 950.

[7] For overviews of this case law, see Durga Das Basu, *Commentary on the Constitution of India* (8th edn, LexisNexis 2011) 59–75; Mohammed Hidayatullah, *Constitutional Law of India* (vol 1, Lucas Publications 1986) 180–231.

[8] Recall that the Irish Constitution simply refers to unconstitutional laws being 'invalid' in Article 15.4, and section 52(1) of the Canadian Constitution Act 1982 requires that unconstitutional law shall be 'without force or effect'.

[9] *Keshavan Madhava Menon v State of Bombay* [1951] INSC 3, [1951] SCR 228.

> [Article 13(1)] provides that all existing laws, insofar as they are inconsistent with the fundamental rights, shall be void to the extent of their inconsistency. *They are not void for all purposes but they are void only to the extent they come into conflict with the fundamental rights*. In other words, on and after the commencement of the Constitution no existing law will be permitted to stand in the way of the exercise of any of the fundamental rights. Therefore, *the voidness of the existing law is limited to the future exercise of the fundamental rights*.[10]

The reasoning in this passage envisages that the law might remain valid in some cases, such as those grounded in pre-constitution events. It might also be considered valid for non-citizens, as only citizens may benefit from the fundamental rights protections in the Indian Constitution. It is important to stress that Das J was here speaking about laws that pre-date the passing of the Indian Constitution. The non-retroactivity point holds up because all it is intended to suggest is that the Indian Constitution does not have retrospective effect with respect to the rights it creates. This is a very positivistic understanding of human rights protection, but a tenable one.

In *Keshavan*, the applicant had commenced their litigation in 1949—prior to the adoption of the Indian Constitution in 1950—and so could not rely on its protections. Insofar as Das J meant to impute that the Constitution could not be used in this case to grant protection to Keshavan, this seems correct. It would have been sufficient merely to hold against Keshavan on this procedural point; it is conspicuously more than this minimal case to assert that Article 13 cannot be read as wiping laws from the statute book altogether. This seems to confuse invalidity with repeal.

The source of this conflation of invalidity and repeal seems to have its source in the judgment of the Bombay High Court in the *Keshavan* case.[11] This has occasionally been reflected directly in statements of the Supreme Court. Take, for example, the way Mahajan CJ characterised the dispute in *Keshavan* in another case three years later:

---

[10] *Keshavan Madhava Menon v State of Bombay* (n 9) 234. Emphasis added.

[11] This is only mentioned in passing in the Supreme Court judgment, but it clearly primed that court in establishing a view that the Court wished to oppose. Mahajan J alludes to this in his judgment: ibid 247.

> The minority view [in *Keshavan*] was that the word 'void' had the same meaning as 'repeal' and therefore a statute which came into [conflict] with fundamental rights stood obliterated from the statute book altogether, and that such a statute was void *ab initio*. The majority however held that the word 'void' in article 13(1), so far as existing laws were concerned, could not be held to obliterate them from the statute book … because in its opinion, article 13 had not been given any retrospective effect. The majority however held that after the coming into force of the Constitution the effect of article 13(1) on such repugnant laws was that it nullified them, and made them ineffectual and nugatory and devoid of any legal force or binding effect. It was further pointed out in one of the judgments representing the majority view, that the American rule that if a statute is repugnant to the Constitution the statute is void from its birth, has no application to cases concerning obligations incurred or rights accrued in accordance with an existing law that was constitutional in its inception, but that if any law was made after the 26th January, 1950, which was repugnant to the Constitution, then the same rule shall have to be followed in India as followed in America. The result therefore of this pronouncement is that the part of the section of an existing law which is unconstitutional is not law, and is null and void.[12]

This is an important passage for a variety of reasons. The equivalence established between voidness and repeal is important. It is also notable that the Court observed that some laws would be 'void' and 'obliterated from the statute book', but that others would not. This lays the groundwork for a non-absolute understanding of voidness. Finally, the reliance on American jurisprudence is notable. As was recounted in chapters 2 and 3, the United States federal courts have a particularly idiosyncratic basis for, and practice of, unconstitutionality. The Indian Constitution, by comparison, is much more prescriptive. Thus, the use of United States federal case law and commentary may be apt to mislead.

6.3.1. Voidness under Article 13(1)

A contrary view to that handed down in *Keshavan* was advanced by some judges in *Behram Khurshed Pesikaka v State of Bombay*.[13] This case concerned certain provisions of the Bombay Prohibition Act 1949, which had been ruled unconstitutional in *State of Bombay v FN Balsara*.[14] The court in *Pesikaka* had to consider the effect of this declaration, a question that promoted divided views. At the first hearing of the

[12] *Behram Khurshed Pesikaka v State of Bombay* [1954] INSC 15, [1955] 1 SCR 613, 651.

[13] ibid.

[14] *State of Bombay v FN Balsara* (n 5).

case, both Venkatarama Ayyar and Jagannadhadas JJ were of the opinion that there was a meaningful distinction to be drawn between a statute that was made outside of legislative competence, and one that was made by a competent legislature, but which infringed some other check on the legislative power.[15]

There are several problems with Venkatarama Ayyar J's approach to this issue. First, his view is partly derived from a comparison with American law.[16] As mentioned above, this is inapt for several reasons. Most obviously, the Indian Constitution provides an incontestable textual basis for judicial review resulting in nullity whereas the US Constitution does not. Furthermore, the Indian Constitution straightforwardly endorses voiding unconstitutional instruments, which are made invalid and inapplicable by operation of Article 13. Unconstitutionality does not, therefore, rely on the operation of precedent barring the application of the law to future claims to achieve widespread effect, as it does in the United States.

More fatally, however, the distinction drawn between breaches of legislative competence and breaches of other checks, such as fundamental rights, itself also seems somewhat capricious.[17] It seems odd to hold that a provincial legislature legislating on a federally-controlled subject should be worse off than a national le-

[15] *Behram Khurshed Pesikaka v State of Bombay* (n 12) 630 (Jagannadhadas J, expressing agreement), 641 (Venkatarama Ayyar J).

[16] In particular, the judge drew on the analysis in Westel Woodbury Willoughby, *The Constitutional Law of the United States* (Baker, Voorhis & Co 1910).

[17] The suggestion that fundamental rights are a check on legislative power and not somehow implicated in the constitution of that power puts the prospective constitutional litigant in the situation of having to say 'it is not that you did not legislate, but rather, you legislated *badly*—so badly, in fact, that the legislation should be disregarded'. This does not fit with the logical structure of constitutional rights. Following a Hohfeldian analysis (see generally: Leif Wenar 'Rights', *The Stanford Encyclopedia of Philosophy* (Fall edn, 2015) ⟨https://plato.stanford.edu/archives/fall2015/entries/rights/⟩ accessed 24th March 2017), constitutional rights can be said to possess several important first- and second-order dimensions. As a first-order concern, I may have a privilege to ϕ (that is, I have no duty not to ϕ) and/or a claim-right that B φ, which in this case could be a negative injunction to B dictating to them not to interfere with the object of my right.
There are also additional, second-order deontic concerns. I have a *power* to dispose of my claim-right, and I also have an *immunity* against others changing the juridical status of my claim-right. My possession of an immunity importantly corresponds to other people possessing a corresponding disability (the opposite or negative of a power). Venkatarama Ayyar J's analysis parses fundamental rights solely as first-order concerns but ignores the second-order matter of the legislature also having a disability concerning interferences with those rights.

gislature legislating on a constitutionally-controlled subject. Both the fundamental rights provisions and the division of powers between the national and state legislatures are creations of the Constitution—so why does one go to competence and one not? Perhaps if one took the view that the national legislature delegates the legislative power to the states then one would be able to hold that *ultra vires* provincial enactments entailed a contradiction in the exercise of that power.[18] This would accommodate review for excess of state jurisdiction. But why not see the Constitution as occupying a similar position vis-à-vis the national legislature? That is to say that the Constitution 'delegates' the legislative power to the national legislature, but with the proviso that the precepts of the whole of the Constitution be obeyed, and not simply some of them.

In any event, this decision was appealed to the Constitution Bench of the court under article 137 of the Constitution.[19] At this second hearing, Mahajan CJ, delivering the majority judgment, expressed disagreement with the distinction drawn by Venkatarama Ayyar J. He reiterated the majority stance in *Keshavan*, to the effect that 'void' in the context of Article 13(1) did not mean 'repeal', as that Article could not be taken to have retrospective effect, reaching into matters that pre-dated the Constitution. In other words, the Constitution could not have authority to repeal laws that pre-existed it. Rather, the voidness in article 13(1) signified that, on the coming into effect of the Constitution, unconstitutional laws are rendered nugatory and of no binding effect and so for the purposes of determining the legal rights and obligations of citizens they should be disregarded.[20]

It is important to note that this logic works for Article 13(1), but not for Article 13(2). Article 13(1) would notionally require retrospective effect to void those pre-constitution laws to which it pertains. Article 13(2), however, makes no claim

[18] In other words, it is a contradiction to be afforded a power that exclusively justifies the performance of certain actions and assert that this same power justifies actions outside those specified.

[19] 'Subject to the provisions of any law made by Parliament or any rules made under article 145, the Supreme Court shall have power to review any judgment pronounced or order made by it'.

[20] *Behram Khurshed Pesikaka v State of Bombay* (n 12) 651–52 (Mahajan CJ).

on any law that is older than itself. For this reason, Mahajan CJ also accepted that, with regard to post-constitution laws, the Indian courts would follow the American rule that unconstitutional statutes be 'void from birth'.[21] Interestingly, just a few years before, Mahajan J (as then he was) adopted a much different American rule in *Dwarkadas Shrinivas of Bombay v The Sholapur Spinning & Weaving Co Ltd*, to the effect that unconstitutional statutes are not void, but rather than official action taken on foot of them is enjoined.[22] In any case, Mahajan CJ's other statements in *Pesikaka* chiefly concern Article 13(1). The Chief Justice also rejected Venkatarama Ayyar J's contention that fundamental rights acted as a check on the exercise of legislative power; rather, he preferred to characterise compliance with fundamental rights as part of the legislative power.[23] Thus, a failure to abide by these norms would constitute a failure of legislative competence.

Das J, in the minority on the second hearing of the case, expressed agreement with Venkatarama Ayyar J's distinction between voidness for want of legislative power and voidness for contradiction of fundamental rights. However, it is notable that even then Das J expressed reluctance to endorse fully the reasoning expressed in his colleague's judgment. In particular, Das J wished to 'guard [himself] against being understood to agree with [Venkatarama Ayyar J's views] relating to waiver of unconstitutionality, the fundamental rights being a mere check on legislative power or the

---

[21] *Behram Khurshed Pesikaka v State of Bombay* (n 12) 651.

[22] *Dwarkadas Shrinivas of Bombay v The Sholapur Spinning & Weaving Co Ltd* [1953] INSC 87 674, 712, citing, with approval, to the rule in *Massachusetts v Mellon* 262 US 447, 448 (1923).

[23] On this point, the judge held that:

> We think that it is not a correct proposition that constitutional provisions in Part III of our Constitution merely operate as a check on the exercise of legislative power. It is axiomatic that when the lawmaking power of a State is restricted by a written fundamental law, then any law enacted and opposed to the fundamental law is in excess of the legislative authority and is thus a nullity. Both these declarations of unconstitutionality go to the root of the power itself and there is no real distinction between them. They represent but two aspects of want of legislative power. The legislative power of Parliament and the State Legislatures as conferred by articles 245 and 246 of the Constitution stands curtailed by the fundamental rights chapter of Constitution. A mere reference to the provisions of article 13(2) and articles 245 and 246 is sufficient to indicate that there is no competency in Parliament or a State Legislature to make a law which comes info clash with Part III of the Constitution after the coming into force of the Constitution.

Behram Khurshed Pesikaka v State of Bombay (n 12) 652 (Mahajan CJ).

effect of the declaration under article 13(1) being "relatively void".'[24] Although he distanced himself from these aspects of Venkatarama Ayyar J's judgment, Das J did support the somewhat contentious proposition that to hold that the invalid part of the legislation was effectively nugatory or destroyed would be 'tantamount to saying, covertly if not openly, that the judicial pronouncement has to that extent amended the section'.[25] In other words, that judicial negation of an unconstitutional provision was tantamount to repeal of that provision.

The subsequent case of *Bhikaji Narain Dhakras v State of MP*[26] again concerned the application of article 13(1), this time to section 43 of the CP and Berar Motor Vehicles (Amendment) Act 1947. In this case, this provision became invalid upon the passing of the Constitution. However, in 1951 a constitutional amendment was passed that would have had the effect of validating the statutory provision.[27] The Court again reiterated that laws found invalid under article 13(1) were not 'dead for all purposes'[28] and that they exist for the purposes of adjudicating pre-constitution fact scenarios and would remain operative even after the passing of the Constitution as against non-citizens. This introduced the 'doctrine of eclipse' into Indian law, which is discussed further below.

In *MPV Sundararamier v State of AP*[29] Venkatarama Ayyar J once again stated his view that laws made in breach of legislative competence were null and void, but laws that breached a fundamental rights provision would be simply unenforceable.[30] The judge also took an opportunity to review both the American and Indian authorities on the subject, and concluded that the correct statement of law was as

---

[24] ibid 670 (Das J).

[25] ibid 668.

[26] *Bhikaji Narain Dhakras v State of Madhya Pradesh* [1955] INSC 48, [1955] 2 SCR 589.

[27] This is not unlike the revival of section 5 of the Punishment of Incest Act 1908 by section 45(3) of the Courts (Supplemental Provisions) Act 1961 in the Irish case of *The People (DPP) v WM* [1995] 1 IR 226 (HC). It is interesting that in the Irish context, a mere legislative enactment (as opposed to a constitutional amendment) sufficed to revive a stricken pre-constitution law.

[28] *Bhikaji Narain Dhakras v State of Madhya Pradesh* (n 26) 600.

[29] *Sundararamier v State of Andhra Pradesh* [1958] INSC 19, [1958] SCR 1422.

[30] The judge's reasoning in more full form was as follows:

follows:

> Where an enactment is unconstitutional in part but valid as to the rest, assuming of course that the two portions are severable, it cannot be held to have been wiped out of the statute book as it admittedly must remain there for the purpose of enforcement of the valid portion thereof, and being on the statute book, even that portion which is unenforceable on the ground that it is unconstitutional will operate *proprio vigore* when the Constitutional bar is removed, and there is no need for a fresh legislation to give effect thereto.[31]

One might be forgiven for thinking that the difference of opinion as to the scope of voidness in Article 13(1) had been resolved by the *Bhikaji* and *Sundararamier* cases. However, in subsequent decisions, the court went on to depart from the *Sundararamier* standard without any reference thereto. So, for example, in *Basheshar Nath vs The Commissioner of Income Tax*[32] Das J again endorsed the distinction between unconstitutionality that went to legislative competence and unconstitutionality that does not, and Subba Rao J rejected it. Similarly, in *Deep Chand*, Subba Rao J restated his objection to the distinction and affirmed Mahajan CJ's original approach in *Keshavan*.[33] Further developments on this distinction are considered below in the discussion on *Jagannath* and *Ambica Mills*.

> Now, in considering the question as to the effect of unconstitutionality of a statute, it is necessary to remember that unconstitutionality might arise either because the law is in respect of a matter not within the competence of the legislature, or because the matter itself being within its competence, its provisions offend some constitutional restrictions. [L]egislation on a topic not within the competence of the legislature and a [sic] legislation within its competence but violative of constitutional limitations have both the same reckoning in a court of law; they are both of them unenforceable. But does it follow from this that both the laws are of the same quality and character, and stand on the same footing for all purposes? This question has been the subject of consideration in numerous decisions in the American Courts, and the preponderance of authority is in favour of the view that while a law on a matter not within the competence of the legislature is a nullity, a law on a topic within its competence but repugnant to the constitutional prohibitions is only unenforceable. … If a law is on a field not within the domain of the legislature, it is absolutely null and void, and a subsequent cession of that field to the legislature will not have the effect of breathing life into what was a still-born piece of legislation and a fresh legislation on the subject would be requisite. But if the law is in respect of a matter assigned to the legislature but its provisions disregard constitutional prohibitions, though the law would be unenforceable by reason of those prohibitions, when once they are removed, the law will become effective without re-enactment.
>
> *Sundararamier v State of Andhra Pradesh* (n 29) 1468–69 (Venkatarama Ayyar J).

[31] ibid 1474–75.

[32] *Basheshar Nath v Commissioner of Income Tax* [1958] INSC 117, [1959] 1 Suppl SCR 528.

[33] *Deep Chand v State of Uttar Pradesh* [1959] INSC 3, [1959] Suppl SCR 8.

6.3.2. Voidness under Article 13(2)

Many of the cases above considered legislation found unconstitutional under article 13(1). At least conceptually, the judiciary's understanding of the effect of article 13(2) was somewhat clearer. For example, in *Saghir Ahmad v State of UP*[34] the UP Road Transport Act 1951, a law enacted after the adoption of the Constitution, was found to be invalid notwithstanding an amendment that would have validated it. This is the same issue as was before the Court in *Bhikaji Narain Dhakras* above; the only difference is that the *Dhakras* case concerned article 13(1) and *Saghir Ahmad* concerned article 13(2). In *Saghir Ahmad* the court reasoned as follows:

> The amendment of the Constitution, which came later, cannot be invoked to validate an earlier legislation which must be regarded as unconstitutional when it was passed: As Professor Cooley has stated in his work on Constitutional Limitations 'a statute void for unconstitutionality is dead and cannot be vitalised by a subsequent amendment of the Constitution removing the constitutional objection but must be re-enacted'. We think that this is sound law and our conclusion is that the legislation in question which violates the fundamental right of the appellants … must be held to be void under article 13(2) of the Constitution.[35]

Thus, laws enacted after the Constitution will be void *ab initio*, but laws enacted prior to it may continue in force against non-citizens. It is therefore surprising that in the course of his judgment in the later *Deep Chand* case, Subba Rao J declared that the word 'void' had the same meaning in both Article 13(1) and Article 13(2).[36] Similarly, in *Mahendra Lal Jaini v State of UP*,[37] Wanchoo J held that the meaning of the word 'void' meant the same thing in both provisions,[38] but that there was a crucial difference that voidness would only be *ab initio* for post-constitution laws.

---

[34] *Saghir Ahmad v State of UP* [1954] INSC 89, [1955] SCR 707.

[35] ibid 728, citing: Thomas Cooley, *A Treatise on the Constitutional Limitations which Rest Upon the Legislative Power of the States of the American Union* (7th edn, first published 1868, Little, Brown & Co 1903) 259, fn 2.

[36] '[A] law, whether pre-Constitution or post-Constitution, would be void and nugatory in so far as it infringed the fundamental rights.' *Deep Chand v State of Uttar Pradesh* (n 33) 40 (Subba Rao J).

[37] *Mahendra Lal Jaini v State of Uttar Pradesh* [1962] INSC 303, [1963] Suppl SCR 912.

[38] He treated the analysis of the word 'void' as something to be taken in the context of Article 13 as a whole, and did not discriminate between the two sub-articles: ibid 934–35.

Pre-constitution laws would become void on the passing of the Constitution and could thus remain valid for certain purposes.

Basu has summarised the effect of Article 13(2) as requiring the following:

> As distinguished from [Article 13(1)], [Article 13(2)] makes the inconsistent laws void *ab initio* and even convictions made under such unconstitutional laws shall have to be set aside. Anything done under the unconstitutional law, whether closed, completed or inchoate, will be wholly illegal and relief in one shape or another has to be given to the person affected by such unconstitutional law. Nor is it revived by any subsequent event.[39]

This is a very stringent effect to attribute to Article 13(2). Importantly, this statement strongly reflects the idea that voidness means different things for pre- and post-constitution law. This is further exemplified by the practice that constitutional amendments cannot save a post-constitution law from unconstitutionality,[40] but they can save a pre-constitution law.[41] However, the strength of this statement must be considered in light of the *Jagannath* and *Ambica Mills* cases, which seem to have placed some qualification on strong endorsements of voidness *ab initio* under Article 13(2).

6.3.3. Some Clarity: *Jagannath* and *Ambica Mills*

In the *Jagannath* case,[42] the Supreme Court seemed to distance itself from the influence of American commentators and, by implication, some of its prior analysis; however, it did not rule definitively on this point. As time passed and more law was created that would fall for consideration under Article 13(2) as opposed to the heretofore more prevalent 13(1), the Supreme Court opined that:

> [A]lthough decisions of the American Supreme Court and the comments of well known commentators like Willoughby and Cooley have great persuasive force, we

---

[39] Durga Das Basu, *Commentary on the Constitution of India* (8th edn, LexisNexis 2011) 63. Internal citations to *Keshavan Madhava Menon v State of Bombay* (n 9); *Deep Chand v State of Uttar Pradesh* (n 33); *Mahendra Lal Jaini v State of Uttar Pradesh* (n 37); *State of Madhya Pradesh v Thakur Bharat Singh* 1967 AIR 1170, [1967] 2 SCR 454.

[40] *Saghir Ahmad v State of UP* (n 34).

[41] *Bhikaji Narain Dhakras v State of Madhya Pradesh* (n 26).

[42] *Jagannath v Authorised Officer, Land Reforms* [1971] INSC 279, [1972] 1 SCR 1055.

> need not interpret our Constitution by too much reliance on them. Nor is it necessary to scrutinise too closely the decisions wherein views appear to have been expressed that a law which is void under Art. 13(2) is to be treated as still-born.[43]

The *Jagannath* case is also interesting as it concerned a post-constitution invalid law, The Madras Land Reforms (Fixation of Ceiling on Land) Act, 1961, which was voided but was held nevertheless to have some kind of existence because it had been put in the Ninth Schedule of the Constitution by a subsequent amendment inserting Article 31-B. This is a highly unusual circumstance, however, and so this bespoke point is unlikely to be of much precedential worth in future cases.

In a landmark treatment of the issue of unconstitutionality—the *Ambica Mills*[44] case—the Court distanced itself from many prior statements and finally seemed to settle on the view that, even for the purposes of article 13(2), voidness was not absolute. The leading judgment was delivered by Mathew J. This judgment is useful for clarifying some confusing points that had arisen in the prior case law. For example, when determining whether 'void' meant the same thing in both Article 13(1) and 13(2), the judge said that '[i]f the meaning of the word 'void' in article 13(1) is the same as its meaning in article 13(2), it is difficult to understand why a pre-Constitution law which takes away or abridges the rights under article 19 should remain operative even after the Constitution came into force as regards non-citizens and a post-Constitution law which takes away or abridges them should not be operative as respects noncitizens'.[45] The judge thus endorsed the view that whatever meaning is to be given to the word void in the context of article 13, it should be univocal in its sub-articles. Of course, this is as-yet silent on whether that concept of voidness is an absolute or relative one.

On this score, Mathew J endorsed voidness as a relative rather than an absolute concept, though some justification was given to this view with respect to particular idiosyncrasies of the Indian constitutional framework:

---

[43] ibid 1069–70.

[44] *State of Gujarat v Shri Ambica Mills* [1974] INSC 72, [1974] 3 SCR 760.

[45] ibid 777.

> [T]the real reason why [an unconstitutional law] remains operative as against non-citizens is that it is void only to the extent of its inconsistency with the rights conferred under Article 19 and that its voidness is, therefore, confined to citizens, as, *ex hypothesi* the law became inconsistent with their fundamental rights alone. If that be so, we see no reason why a post-Constitution law which takes away or abridges the rights conferred by article 19 should not be operative in regard to non-citizens as it is void only to the extent of the contravention of the rights conferred on citizens.[46]

Article 13 is thus most plausibly understood as regulating the applicability of unconstitutional law with respect to citizens and non-citizens. This is justifiable on the basis that Part III of the Indian Constitution only confers fundamental rights on *citizens*. As such, it does not necessarily require a broader legal validity claim. There is further evidence of this shying away from a legal validity claim when the judge later claims that:

> [V]oidness is not *in rem* but to the extent only of inconsistency or contravention, as the case may be of the rights conferred under Part III. Therefore, when article 13(2) uses the expression 'void', it can only mean, void as against persons whose fundamental rights are taken away or abridged by a law. The law might be 'still-born' so far as the persons, entities or denominations whose fundamental rights are taken away or abridged, but there is no reason why the law should be void or 'still-born' as against those who have no fundamental rights.[47]

Although this is perfectly tenable as a claim about legal applicability, it is less clear how voidness could be anything but *in rem*.[48] It is not relative to context and circumstance. It seems best, therefore, to view *Ambica Mills* as representing a more concrete shift on the part of the Indian judiciary towards an applicability-centred view of judicial review for unconstitutionality, rather than one that focuses on more absolute and binary concepts, such as voidness, validity or existence of the law.

Mathew J also took the opportunity to review the still-unresolved dispute in the prior case law between Das J/Venkatarama Ayyar J and Mahajan CJ/Subba Rao J on the topic of whether a breach of fundamental rights went to legislative competence or not. On this point he thought that there was 'nothing strange in the notion of a le-

---

[46] *State of Gujarat v Shri Ambica Mills* (n 44) 777.

[47] ibid 777–78.

[48] I discuss in chapters 8, 9 and 10 how legal validity is relative in many ways. One of the ways in which it is relative is that it must be related to the *existence* (system membership) of other laws.

gislature having no inherent legislative capacity or power to take away or abridge by a law the fundamental rights conferred on citizens and yet having legislative power to pass the same law in respect of noncitizens who have no such fundamental rights to be taken away or abridged. In other words, the legislative incapacity … in this context would be the taking away or abridging by law the fundamental rights under Article 19 of citizens'.[49] In later cases, this point seems to have been resolved in favour of the view that there is no meaningful difference here.[50]

In summary, although the Indian authorities were initially somewhat confusing on the issue of unconstitutionality, and they are occasionally poorly anchored to their Constitutional text, it seems best to conclude that India suspends or enjoins the enforcement of unconstitutional statutes rather than voiding them.[51] In this respect, it is similar to the United States. Nor is this similarity accidental; indeed, it seems to be precisely because of an endorsement of a faulty distinction between types of legislative failure, and a desire to emulate United States jurisprudence. Perhaps the most useful lesson to be drawn from the general Indian experience for this thesis is that constitutional text may fail to substantially guide practice. This suggests that theoretical and practical concerns may influence unconstitutionality even notwithstanding a constitutional text.

6.4. THE DOCTRINE OF ECLIPSE

Through the case law on Article 13 discussed above, the Indian courts developed a 'doctrine of eclipse'. This doctrine applies specifically to laws that are invalidated on the basis of the Part III fundamental rights provisions, distinguishing between pre- and post-Constitution laws. We have seen already that, particularly with respect to

[49] *State of Gujarat v Shri Ambica Mills* (n 44) 779.

[50] 'When a statute is in part void, it will be enforced as regards the rest, if that is severable from what is invalid. It is immaterial for the purpose of this rule whether the invalidity of the statute arises by reason of its subject-matter being outside the competence of the legislature or by reason of its provisions contravening constitutional prohibitions'. *Ashoka Kumar Thakur v Union of India* [2008] INSC 614 [23] (Arijit Pasayat J).

[51] The doctrine of eclipse, considered immediately below, would also tend to support this view.

Article 13(1), the word 'void' does not mean negated entirely for the purposes of Indian judicial review of legislation. As the name of the doctrine under discussion would suggest, legislation impugned under Article 13(1) simply becomes dormant or 'eclipsed'. Continuing with this metaphor, if the obstructing object causing the eclipse is removed, then the legislation can become reanimated.

The Supreme Court introduced this doctrine to India's constitutional jurisprudence in *Bhikaji Narain Dhakras v State of Madhya Pradesh*.[52] This case concerned the pre-constitution CP and Berar Motor Vehicles Amendment Act 1947, which had been struck down. However, a subsequent amendment to the Constitution cured the infirmity and made the legislation compatible with the Constitution again. The Court held that the legislation was given new life by this process. Thus, the net effect of the doctrine of eclipse is to achieve a suspension of the applicability of a provision. Das CJ, delivering the judgment for the court, observed that:

> All laws, existing or future, which are inconsistent with the provisions of Part III of our Constitution are by the express provision of article 13, rendered void 'to the extent of such inconsistency'. Such laws were not dead for all purposes. They existed for the purposes of pre-Constitution rights and liabilities and they remained operative, even after the Constitution, as against non-citizens. It is only as against the citizens that they remained in a dormant or moribund condition.[53]

As is clear from the above quotation, the original formulation of the doctrine in the *Dhakras* case only applied to pre-constitution laws. This mirrors the distinction in Ireland, as seen in chapter 5, between invalidity under Article 15 and unenforceability under Article 50. This approach was repeated in *Purshottam Govindji Halai v Shree BM Desai* where again Das CJ declared that 'The effect of article 13(1) is that the law could not stand in the way of the enjoyment of fundamental rights. *The law was not dead*'.[54]

These initial cases leave the doctrine's application to post-Constitution laws as

[52] *Bhikaji Narain Dhakras v State of Madhya Pradesh* (n 26).

[53] ibid 599–600 (Das CJ).

[54] *Purshottam Govindji Halai v Shree BM Desai* [1955] INSC 53, [1995] 2 SCR 887, 904 (Das CJ). Emphasis added.

an open question. In both *Deep Chand*[55] and *Mahendra Lal Jaini*[56] the Supreme Court expressed some doubt as to whether the doctrine could be extended to post-Constitution laws. In *Deep Chand*, Subba Rao J, delivering the majority judgment, simply declared that '[t]here is no scope for applying the doctrine of eclipse to a case where the law is void *ab initio* in whole or in part'.[57] Given that laws invalidated under Article 13(2) are taken to be void *ab initio*, this seems to rule out the application of the doctrine of eclipse to post-Constitution laws. Das CJ, the original author of the doctrine in the *Bhikaji* case, dissented and rejected the view that the doctrine of eclipse could not apply to any post-Constitution law.[58] Once again, he focused on the case of the non-citizen as a potential example:

> A post-Constitution law may infringe either a fundamental right conferred on citizens only or a fundamental right conferred on any person, citizen or non-citizen. In the first case the law will not stand in the way of the exercise by the citizens of that fundamental right and, therefore, will not have any operation on the rights of the citizens, but it will be quite effective as regards non-citizens. In such a case the fundamental right will … throw a shadow on the law which will nevertheless be on the Statute Book as a valid law binding on non-citizens and if the shadow is removed by a constitutional amendment, the law will immediately be applicable even to the citizens without being re-enacted.[59]

In the later *Mahendra Lal Jaini* case, however, Das CJ's view did not find support. Wanchoo J delivering judgment for the court found, after reviewing the authorities, that 'the doctrine of eclipse will apply to pre-Constitution laws which are governed by Art. 13(1) and would not apply to post-Constitution laws which are governed by Art. 13(2)'.[60] There is, thus, a degree of inconsistency as to how far the doctrine is thought to go. Inconsistency that, as Chandrachud has observed in a recent brief survey of the doctrine,[61] has not completely dissipated.

---

[55] *Deep Chand v State of Uttar Pradesh* (n 33).

[56] *Mahendra Lal Jaini v State of Uttar Pradesh* (n 37).

[57] *Deep Chand v State of Uttar Pradesh* (n 33) 38.

[58] ibid 13.

[59] ibid 13.

[60] *Mahendra Lal Jaini v State of Uttar Pradesh* (n 37) 935.

[61] Chintan Chandrachud, 'Declarations of Unconstitutionality in India and the UK: Comparing the

6.4.1. Extension of the Doctrine to Post-Constitution Law?

It is still unclear whether the doctrine of eclipse can be extended to post-constitution law. In one of the most recent considerations of the doctrine of eclipse—*KK Poonacha v State of Karnataka*—the Court, after reviewing past authorities, held:

> [A] post-Constitution law is void *ab initio* if it is not within the domain of the Legislature or is violative of the rights conferred by Part III of the Constitution. If the law is within the legislative competence of the Union or State and does not infringe any of the rights conferred by Part III of the Constitution, then the same cannot be declared void on the ground of non compliance of the procedural requirement of prior recommendation or sanction, if assent is given in the manner provided under Article 255 of the Constitution. … [I]f a law is within the competence of the Legislature, the same does not become void … merely because post enactment assent of the President has not been obtained. Such law remains on the statute book but cannot be enforced till the assent is given by the President. Once the assent is given, the law becomes effective and enforceable. If the provision requiring pre enactment sanction or post enactment assent of the President is repealed, then the law becomes effective and enforceable from the date of repeal and such law cannot be declared unconstitutional only on the ground that the same was not reserved for consideration of the President and did not receive his assent.[62]

The material point here is that the understanding of voidness as relative for post-constitution law seems to be resurgent in the *Poonacha* judgment. At least so far as assent procedures are concerned, the Court's judgment seems to envisage that voidness can be 'cured' through subsequent assent or removal of other impediments to commencement. This is so notwithstanding the theory that unconstitutional law is a nullity and so, notionally, there should be nothing to commence. If voidness is relative for post-constitution law, this gives greater scope to extent the doctrine of eclipse to apply to such law. The continuing influence of United States jurisprudence in the evolution of this aspect of Indian law is also evident here. Singhvi J began his analysis with a quote from Willoughby saying 'The Court does not annul or repeal the statute if it finds it in conflict with the Constitution. It simply refuses to recognize it, and determines the rights of the parties just as if such statute had no application'.[63] As I argued above when criticising the view of Venkatarama Ayyar J in *Pesikaka*

Space for Political Response' (2015) 43 Ga J Int'l & Comp L 309, 330–31.

[62] *KK Poonacha v State of Karnataka* [2010] INSC 706 [20].

[63] ibid [8], citing: Willoughby (n 16).

and *Sundararamier*, the United States Constitution and constitutional praxis are not a particularly apposite model for India. Since, at the time of *Poonacha*, there was ample comparison that could have been drawn with closer comparator jurisdictions (Canada and Ireland, to name but two), it is somewhat disappointing to see that the Indian Courts continue to draw on older American authorities. However, in light of *Ambica Mills* endorsing a more 'relative' understanding of voidness generally, these authorities might be more appropriate now than they were before, notwithstanding the difficulty of reconciling them with a coherent concept of voidness.

The judgment in *Poonacha* can also be compared with the slightly earlier *Vij* case, which was not cited in the *Poonacha* decision, where the court found that 'having regard to the prohibition contained in clause (2) of Article 13 of the Constitution any law made in contravention of Part III of the Constitution would be a stillborn law and such a law is dead from the very beginning. A law, which is stillborn and is dead right from its inception, cannot at all be taken notice of or read for any purpose whatsoever'.[64] This reasoning would seem to block the extension of the doctrine of eclipse to post-constitution case law.[65]

6.4.2. Restrictions on the Doctrine of Eclipse

Some restrictions on the doctrine have been made apparent in the case law. Currently, it applies only to claims of unconstitutionality under Article 13(1) of the Indian Constitution.[66] As has been observed above, some Supreme Court precedents support the view that the doctrine of eclipse can operate to restore the applicab-

[64] *Rakesh Vij v Raminder Pal Singh Sethi* [2005] INSC 522 [11].

[65] It is interesting to compare this with how the Irish legislature was apparently able to take notice of section 44(2) of the Road Traffic Act 1968 by amending it in section 7 of the Road Traffic (Amendment) Act 1973, notwithstanding that section 44(2) of the 1968 Act had been found unconstitutional (and thus void *ab initio*) in *Maher v Attorney General* [1973] IR 140 (SC).

[66] *State of Rajasthan v Hindustan Sugar Mills Ltd* [1988] INSC 174, [1988] 1 Suppl SCR 461, 469.

ility of post-1950 invalid legislation.[67] Others cut in the opposite direction.[68] The High Courts have also followed the Supreme Court's lead on this point, with some favouring the view that the doctrine of eclipse can operate on post-Constitution legislation[69] and others disfavouring the same view.[70]

Squaring the doctrine with the strict language of Article 13 is already a difficult exercise, even more so with Article 13(2). It seems best to consider it as operating similarly to the Irish Constitution's Article 50, which recognises that laws validly passed by previous legislative authorities are still 'valid' laws (albeit ones that will be unenforceable in the Republic of Ireland if found unconstitutional there). There are statements of the Indian judiciary that would lend support to this analysis. In distinguishing why the doctrine of eclipse might apply to pre-1950 laws, but not post-1950 laws, the Court has held that:

> The American authorities refer only to post-Constitution laws which were inconsistent with the provisions of the Constitution. Such laws never came to life but were still-born as it were. The American authorities, therefore, cannot fully apply to pre-Constitution laws which were perfectly valid before the Constitution.[71]

It is helpful to consider Articles 13(1) and 13(2) separately here. The 'American' view in the Indian courts is the view that unconstitutionality means invalidity and voidness *ab initio*. This view supposedly attaches to unconstitutional law under Article 13(2). Article 13(1), on the other hand, seems to preclude the Indian Courts from adjudicating on legal validity, on a source-based conception of legal validity. It assumes that the pre-constitution law is valid in some other past system, but not the current legal system. However, in the same way as uncommenced legislation, un-

---

[67] *Sundararamier v State of Andhra Pradesh* (n 29); *State of Gujarat v Shri Ambica Mills* (n 44); *KK Poonacha v State of Karnataka* (n 62).

[68] *Deep Chand v State of Uttar Pradesh* (n 33); *Mahendra Lal Jaini v State of Uttar Pradesh* (n 37); *Rakesh Vij v Raminder Pal Singh Sethi* (n 64).

[69] *PL Mehara v DR Khanna* AIR 1971 Delhi 1; *Minoo Framroze Balsara v Union of India* AIR 1992 Bom 375.

[70] *Chembakave Vadakkekkara Lakshmi v Nellisseri Gramam Narayanaswami* AIR 1963 Ker 330; *Bawa Singh v Union of India* 6 (1970) DLT 409; *Ram Chand v State of Haryana* (1971) 73 PLR 958; *Dharam Pal v Kaushalya Devi* AIR 1990 Raj 135.

[71] *Bhikaji Narain Dhakras v State of Madhya Pradesh* (n 26) 599.

constitutional pre-constitution law may not have the force of law in India (though this may be cured in a variety of ways, hence the 'doctrine of eclipse'). This seems to require a different understanding of voidness in both provisions: voidness is 'relative' in Article 13(1), but 'absolute' in Article 13(2).

The view of voidness under Article 13(2) as 'relative' in *Ambica Mills* (but not in *Vij*) could allow for the doctrine of eclipse to extend to post-constitution law as well. However, this seems to be difficult to accommodate under the text of Article 13(2) itself. In many respects, the debate over the extension of the doctrine of eclipse to post-constitution law is a symptom of a much deeper and more difficult tension between the constitutional text on one hand, and the endorsement of conflicting understandings of voidness on the other. This difficulty does not seem to have yet been resolved by the Indian courts.

6.5. CONCLUSION

In some respects, the Indian position on unconstitutionality can be quite confusing. There are many inconsistent statements on the nature of unconstitutionality. In many cases there was little agreement on whether unconstitutionality is a failure of the legislative power or a positive intervention by the judicial power. This confusion is compounded against the background of a constitutional text that seems to, quite restrictively, require that unconstitutional legislation be simply found 'void'. It is small wonder that the Indian courts have navigated around this stricture; however, the way in which this has been achieved has been *ad hoc*, and it is difficult to discern significant trends from the case law that reflect consistent and coherent theories of unconstitutionality. The *Ambica Mills* decision is to be praised for bringing some clarity to this area of Indian jurisprudence. However, in many respects, that decision confirms a theory of unconstitutionality that is poorly fitted to the Indian Constitutional text, and this tension emerges in the inconsistency even between more recent cases, such as *Vij* and *Poonacha*.

Much of this difficulty seems to have been grounded in the Indian judiciary's ap-

plication of United States precedents and thought to the issue of unconstitutionality without squaring this thought with their constitutional text. Analysis is made more difficult here because, as demonstrated in chapters 2 and 3, there are inconsistent statements from United States federal courts as to the effect of unconstitutionality. A more general lesson that can be learned from this is the danger of the application of foreign law in cases such as these. This is a theme that I return to in chapter 11.

The findings of this thesis could be used to bring greater clarity to Indian practice. Consider the difficulty presented by the divergent meanings of 'void' for Article 13(1) and 13(2). A validity-centred analysis is likely the reason voidness differs in meaning between these provisions. The Indian courts consider that they cannot pronounce on the validity of pre-constitution law, but (although this was subject to some controversy in the case law) the Indian legislature is not competent to pass unconstitutional law. This seems to result in a situation where pre-constitution law may be valid (by reference to some past criterion of validity) but suspended in application, but post-constitution law may be invalid. The link between absolute voidness and invalidity is predicated on the assumption that invalid law is *non-existent* and *inapplicable*. This is an assumption that I challenge in chapter 8. In the absence of this assumption, it would be possible in principle to endorse the analysis of relative voidness in *Ambica Mills* and to fit this analysis to the Indian constitutional text.

7 | South Africa

CHAPTER OVERVIEW

7.1. Introduction 165

7.2. Doctrines Avoiding or Curtailing Unconstitutionality 166

7.2.1. Reading In 166

7.2.2. Severance 168

7.2.2.1. Actual Severance 169

7.2.2.2. Notional Severance 170

7.3. The Doctrine of 'Objective Invalidity' 171

7.3.1. The Basis of the Doctrine: *Ferreira v Levin* 172

7.3.2. Recent Applications of the Doctrine 174

7.3.3. *Erga Omnes* Effect 177

7.4. Suspended Declarations of Unconstitutionality 178

7.4.1. Constitutional Basis and Frequency of Suspension 179

7.4.2. Interim Orders 181

7.5. Conclusion 182

7.1. INTRODUCTION

The South African constitution is the youngest constitution considered in this thesis and, in many respects, the most innovative. Specifically on the issue of unconstitutionality, it is the only constitution with a textually-endorsed regime of suspending declarations of unconstitutionality. Given this textually legitimate basis, it is interesting to compare how the remedy has been used here as against its use in Canada, where it was a judicial innovation. The Court also has a wide power to impose any

order that is 'just and equitable' (section 172 of the Constitution) to remedy an instance of unconstitutionality.[1]

Regarding the meaning of unconstitutionality in the South African Courts, as was seen in chapter 2 there is unambiguous support textual support for the proposition that unconstitutionality means invalidity. This is further reflected in the courts' jurisprudence on the doctrine of 'objective invalidity', considered in this chapter.

While the South African constitution does also provide for the continuance in force of pre-constitution law in Schedule 6, Item 2, such law is not treated materially differently from law enacted after the passing of the Constitution. This is accounted for below in the discussion of the doctrine of 'objective invalidity'. As such, pre-constitution law is not given separate consideration in this chapter.

7.2. DOCTRINES AVOIDING OR CURTAILING UNCONSTITUTIONALITY

7.2.1. Reading In

The South African Constitutional Court has minimised the disparity between adding and deleting words, and has said that there is little in the difference between reading in and severance. They have acknowledged that they mirror each other in that one changes legislation by excision and one by addition.[2] The court has also drawn attention to the acute separation of powers concerns underlying these remedies, suggesting that reading in and severance are both justified as they are non-final:

---

[1] The meaning of a 'just and equitable' order in this context was considered recently by the Constitutional Court:

> That may include an order: (a) that is prospective; (b) whose retrospectivity is not limited and is thus—in accordance with the principle of 'objective invalidity'—retrospective to the date the Constitution, in the case of pre-Constitution legislation, or the legislation in issue, in the case of post-Constitution legislation, came into effect; (c) that limits the retrospective effect of the declaration of invalidity; or (d) that suspends the declaration of invalidity—with or without interim relief during the period of suspension—to afford the Legislature an opportunity to remedy the defect.

Ramuhovhi v President of the Republic of South Africa [2017] ZACC 41, (2018) 2 SA 1 [44]

[2] *National Coalition for Gay and Lesbian Equality v Minister of Home Affairs* [1999] ZACC 17, (2000) 2 SA 1 [67].

> [W]hether the remedy a court grants is one striking down, wholly or in part; or reading into or extending the text, its choice is not final. Legislatures are able, within constitutional limits, to amend the remedy, whether by re-enacting equal benefits, further extending benefits, reducing them, amending them, 'fine-tuning' them or abolishing them. Thus, they can exercise final control over the nature and extent of the benefits.[3]

Regarding the factors that the court thought would be relevant in guiding future courts in deciding to sever from, or read words into, a statute, the Court adopted a two-step test. First, a court would have to ensure that whatever the provision that would result from the severance or reading in is itself consistent with the Constitution. Second, a court must ensure that the resultant provision would minimally interfere with other legislative provisions.[4] Concerning reading in specifically, a court must also be able to define the proposed extension to the law precisely, and it must be as faithful as possible to the legislative scheme into which it is reading in.[5] Although the Court in *National Coalition for Gay and Lesbian Equality* also suggested that it would not grant the remedy of reading in where it would result in an 'unsupportable budgetary intrusion',[6] this does not seem to have been followed in subsequent cases where the remedy was nevertheless granted despite significant budgetary impact.[7]

More recently, the Court has seemingly retreated from the view that it should lay down much guidance in advance, preferring the view that the appropriateness of particular remedies will be determined on a case by case basis:

> How a court exercises its duties to remedy the constitutional invalidity of a statute calls for a degree of restraint in appropriate circumstances, The extent to which a court should refrain from interfering in the legislative realm, however, will largely be determined by the facts and circumstances of each case, for which reason it would be undesirable to lay down a general rule as to when or how a court should do so.[8]

---

3 ibid [76].

4 ibid [74].

5 ibid [75].

6 ibid [75].

7 *Khosa v Minister of Social Development; Mahlaule v Minister of Social Development* [2004] ZACC 11, (2004) 6 SA 505 [88].

8 *C v Department of Health and Social Development, Gauteng* [2012] ZACC 1, (2012) 2 SA 208 [44].

The relative laxity in imposing a strict standard here may however be partly explained, or at least ameliorated, by the Court's consistent statement that it is open for the legislature to amend a reading in where they disagree with or disapprove of the court's imposed solution.[9]

7.2.2. Severance

As the language of section 172(1)(a) only requires the courts to declare a law invalid *to the extent of its inconsistency with the Constitution*, there is clear textual scope for severing unconstitutional parts of laws and thus quarantining the effect of a full-blown declaration of unconstitutionality. The South African courts recognise two types of severance: *actual* severance, and *notional* severance. The first of these is familiar and operates similarly to severance doctrines in other jurisdictions: where possible, unconstitutional phrasing is excised, and the rest of the section is left intact. The second is somewhat more idiosyncratic and is discussed in greater detail below.

The Constitutional Court has been proactive in using remedies to mitigate what may have been otherwise severe effects of a declaration. For example, in *Jaftha v Schoeman*, the Court was quick to observe that it was compelled to render constitutionally inconsistent legislation invalid to the extent of that inconsistency,[10] but it immediately limited the scope of that finding by refusing to strike the legislation down. Instead, it found that there would only be one way to remedy the constitutional deficiency here (a lack of judicial oversight) and rectified this deficit by inserting words into the impugned statute that would give it a constitutional reading.[11]

---

[9] *Gaertner v Minister of Finance* [2013] ZACC 38, (2014) 1 SA 442 [84]; *Provincial Minister for Local Government, Environmental Affairs and Development Planning, Western Cape v Municipal Council of the Oudtshoorn Municipality* [2015] ZACC 24, (2015) 6 SA 115 [29].

[10] *Jaftha v Schoeman; Van Rooyen v Stoltz* [2004] ZACC 25, (2005) 2 SA 140 [61].

[11] ibid [63]–[64].

7.2.2.1. Actual Severance

The test for actual severance employed in South African courts was laid down in the joined cases of *Coetzee and Matiso*, and it is to ask whether 'the good is not dependent on the bad'.[12] This test is administered in two steps. First the court asks 'is it possible to sever the invalid provisions' and second 'if so, is what remains giving effect to the purpose of the legislative scheme'[13] In later cases, the court has also stressed that the prior condition of an *inconsistency* with the Constitution (as opposed to the mere desirability of a change to the legislation) must be made out before severance can be performed.[14]

The Constitutional Court has tended to treat the remedy of actual severance with a degree of caution, and it is keenly aware of the danger of judicial rewriting of legislation. In the joined cases of *Case and Curtis*, the Constitutional Court reflected on the experience of the Canadian and Indian Courts[15] and cautioned against injudicious use of severance:

> I do not think that the severance of one or two isolated words … within the challenged definition is a viable option. That is because the offending overbreadth cannot be laid at the door of any one word, or group of words, but rather permeates the entire text. … [I]f we apply a blue pencil to each and every noun form and transitive verb that presents overbreadth problems, we effectively write a new provision that bears only accidental resemblance to that enacted by Parliament. … For this Court to attempt that textual surgery would entail it departing fundamentally from its assigned role under our Constitution. It is trite but true that our role is to review, rather than to re-draft, legislation.[16]

---

[12] *Coetzee v Government of the Republic of South Africa; Matiso v Commanding Officer Port Elizabeth Prison* [1995] ZACC 7, (1995) 4 SA 631 [16.] The test continues to be applied in Constitutional Court cases, for example: *Oriani-Ambrosini, MP v Sisulu, MP Speaker of the National Assembly* [2012] ZACC 27, (2012) 6 SA 588 [93]; *Teddy Bear Clinic for Abused Children v Minister of Justice and Constitutional Development* [2013] ZACC 35, (2014) 2 SA 168 [106]; *Sarrahwitz v Martiz NO* [2015] ZACC 14, (2015) 4 SA 491 [72].

[13] *Coetzee v Government of the Republic of South Africa; Matiso v Commanding Officer Port Elizabeth Prison* [1995] ZACC 7, (1995) 4 SA 631 [16].

[14] *South African Transport and Allied Workers Union v Garvas* [2012] ZACC 13, (2013) 1 SA 83 [131].

[15] Citing *Schachter v Canada* [1992] 2 SCR 679 (SCC) and *R M D Chamarbaugwala v Union of India* [1957] INSC 32, [1957] SCR 930: *Case v Minister of Safety and Security; Curtis v Minister of Safety and Security* [1996] ZACC 7, (1996) 3 SA 617 [74], [fn]109.

[16] ibid [71]–[73].

The references to the Canadian and Indian experiences here show that severance reflects similar concerns in different jurisdictions and has been handled relatively uniformly by the courts in the jurisdictions under study in this thesis.

7.2.2.2. Notional Severance

The principal difference between actual and notional severance is that when notional severance is engaged, the unconstitutional words are not excised from the impugned section; they are, rather, given a particular meaning, or come with the caveat that they can only be employed in certain contexts. Thus, notional severance enables courts to require other organs that interpret and apply legislation to do so in a particular way.

The primary use of notional severance is in situations where the deletion of the words, or the utilisation of an alternative remedy, would leave a gap in the law. In *Islamic Unity Convention v Independent Broadcasting Authority*, the Constitutional Court had to assess the constitutionality of a provision that prohibited the broadcasting of material 'likely to prejudice relations between sections of the population'. The Court found this proscription to be an unconstitutionally overbroad infringement on free speech. In considering its remedial options, however, the Court observed that if the relevant portion were struck down 'in its entirety with nothing to replace it, a dangerous gap would result'.[17] Of particular concern was that the gap would arise in an area where the Constitution mandated regulation; section 192 of the Constitution requires the state to establish an independent broadcasting authority. Ultimately, notional severance was considered apposite in this case as this served the dual aim of making the unconstitutional element of the clause inoperable but allowed the court to require the legislation to be read as including a prohibition preventing the broadcasting of unprotected expression (such as hate speech) thus maintaining a minimal level of regulation in the area.[18]

---

[17] *Islamic Unity Convention v Independent Broadcasting Authority* [2002] ZACC 3, (2002) 4 SA 294 [54].

[18] ibid [55].

A further proviso on the use of notional severance is that it cannot be used where the invalidity of a provision results from an omission. It thus only applies to things that are positively legislated. In *National Coalition for Gay and Lesbian Equality v Minister of Home Affairs* the Constitutional Court held that:

> Where … the invalidity of a statutory provision results from an omission, it is not possible … to achieve notional severance by using words such as 'invalid to the extent that', or other expressions indicating notional severance. An omission cannot, notionally, be cured by severance. … The only logical equivalent to severance, in the case of invalidity caused by omission, is the device of reading in.[19]

The Court observed that in neither the United States nor Canada was there a practice of using severance to cure unconstitutionality resulting from omissions. However, the court alluded to cases in both jurisdictions that saw the reading in of words to a statute where it was unconstitutionally under-inclusive.[20]

7.3. THE DOCTRINE OF 'OBJECTIVE INVALIDITY'

The South African courts consider unconstitutionality to entail invalidity and voidness *ab initio*. This is reflected in the language of the Constitutional Court in *Kruger v President of the Republic of South Africa*, wherein the Court found Proclamation R27 of 2006—purporting to bring into operation sections 4, 6, 10, 11 and 12 of the Road Accident Fund Amendment Act 2005—'null and void and of no force and effect'.[21] Further, the Court found that the President could not amend the proclamation made precisely *because* it was void: 'the President does not have the power to amend a proclamation issued in error where the original proclamation was void from its commencement, as in this case. I cannot see that a nullity can be amended.'[22]

---

[19] *National Coalition for Gay and Lesbian Equality v Minister of Home Affairs* [1999] ZACC 17, (2000) 2 SA 1 [64].

[20] With respect to this practice in the United States, the Court turned its attention to *Iowa-Des Moines Nat Bank v Bennett* 284 US 239 (1931); *Skinner v Oklahoma* 316 US 535 (1942); *Welsh v United States* 398 US 333 (1970); *Califano v Westcott* 443 US 76 (1979).
Regarding Canada, it relied mainly on the authority of *Miron v Trudel* [1995] 2 SCR 418 (SCC) [177]–[181].

[21] *Kruger v President of the Republic of South Africa* [2008] ZACC 17, (2009) 1 SA 417 [80].

[22] ibid [61].

Although these remarks were made in the context of an executive ordinance, the understanding of these concepts transfers soundly to the context of judicial review of legislative action. This also chimes with comments regarding the amendment of nullities observed in chapters 5[23] and 6.[24]

7.3.1. The Basis of the Doctrine: *Ferreira v Levin*

The most stark confirmation of the voidness *ab initio* theory in the South African courts is the doctrine of 'objective invalidity'. This doctrine was first expounded in *Ferreira v Levin*.[25] This case concerned the constitutionality of certain provisions of the Companies Act 1973 that dealt with the examination of persons in winding-up proceedings. The issue was whether this provision was consistent with the privilege against self-incrimination, as statements made in the context of the examination could be held against the individual in subsequent criminal proceedings. The section was declared invalid to the extent that this subsequent admission of evidence would be unconstitutional.[26] In the course of considering the effect of unconstitutionality, and the role of the judicial power in unconstitutionality, the Court found that 'the Court's order does not invalidate the law; it merely declares it to be invalid'.[27] Thus, the court plays no constitutive role in 'creating' the invalidity; it merely recognises, accepts, and declares something that was already the case.[28] Furthermore, it seemed to roundly reject an applicability-based analysis, in the style of the United States federal courts or Indian 'relative' voidness, which it characterised as a 'subjective' approach:

> A statute is either valid or 'of no force and effect to the extent of its inconsistency'. The subjective positions in which parties to a dispute may find themselves cannot have a

[23] *Brennan v Attorney General* [1983] ILRM 449 (HC) 479.

[24] *Saghir Ahmad v State of UP* [1954] INSC 89, [1955] SCR 707.

[25] *Ferreira v Levin* [1995] ZACC 13, (1996) 1 SA 984 [25]–[30].

[26] ibid [156], [157].

[27] ibid [27].

[28] See also: *Fose v Minister of Safety and Security* [1997] ZACC 6, (1997) 3 SA 786 [94].

> bearing on the status of the provisions of a statute under attack. The Constitutional Court, or any other competent Court for that matter, ought not to restrict its enquiry to the position of one of the parties to a dispute in order to determine the validity of a law. The consequence of such a (subjective) approach would be to recognise the validity of a statute in respect of one litigant, only to deny it to another. Besides resulting in a denial of equal protection of the law, considerations of legal certainty, being a central consideration in a constitutional state, militate against the adoption of the subjective approach.[29]

The court has expounded further on this doctrine in numerous decisions, and it is a firmly ingrained aspect of South African constitutional law.[30] It also continues to find favour in more recent decisions of lower appellate courts. The Supreme Court of Appeal has encapsulated the effect and rationale of the doctrine as follows:

> It is clear that a pre-existing provision of a law which is unconstitutional became invalid at the moment the Constitution took effect. This is the effect of the supremacy clause of the Constitution (s 2), in terms of which the Constitution is the supreme law of the Republic and all law or conduct inconsistent with it is invalid. Item 2(1) of Sch 6 to the Constitution provides that all law that was in force when the Constitution took effect, continues in force until amended or repealed, but only to the extent that it is consistent with the Constitution. In accordance with the doctrine known as 'objective constitutional invalidity', a court making a declaration of invalidity simply declares invalid what has already been invalidated by the Constitution.[31]

It is interesting to observe that the above quotation from *Geldenhuys* envisages that the doctrine of objective invalidity applies to pre-constitution law. *Ferreira* itself and more recent applications of the doctrine confirm that it also applies to post-constitution law. For example, in *Shoprite Checkers* the Constitutional Court considered that an Act passed in 2003 was objectively invalid.[32] As shall be seen below, it is also clear that the doctrine applies to administrative actions. The Court does not, therefore, seem to categorise its practice based on the context of the invalidity; rather, it has one concept of invalidity, and it applies that concept in a variety

---

[29] *Ferreira v Levin* (n 25) [26].

[30] *Shoprite Checkers (Pty) Limited v Member of the Executive Council for Economic Development* [2015] ZACC 23, (2015) 6 SA 125 [31]; *Gory v Kolver NO* [2006] ZACC 20, (2007) 4 SA 97 [39]; *Ingledew v Financial Services Board* [2003] ZACC 8, (2003) 4 SA 584 [20]; *Moise v Greater Germiston Transitional Local Council* [2001] ZACC 2, (2001) 4 SA 1288 [12]–[14]; *Prince v President, Cape Law Society* [2000] ZACC 28, (2001) 2 SA 388 [36]–[37].

[31] *Geldenhuys v State* [2008] ZASCA 47, (2009) 1 LRC 294 [76].

[32] *Shoprite Checkers (Pty) Limited v Member of the Executive Council for Economic Development* [2015] ZACC 23, (2015) 6 SA 125 [31].

of contexts. The logic is that the Constitution, when passed, immediately invalidated all unconstitutional laws. The cases establishing this (notably, *Ferreira*) refer to the Interim Constitution, but the Final Constitution makes similar provisions. Section 2 states that law inconsistent with the Final Constitution *is* invalid.[33] Section 172(1)(a) positively requires a court to *declare* unconstitutional law invalid.[34]

As potentially severe as the doctrine of objective invalidity might sound on its face, the South African courts have developed a sophisticated practice and suite of approaches that help to mollify the effects of unconstitutionality by either avoiding a declaration of unconstitutionality where possible or by limiting the scope of such declarations where they must be made. These practices were considered above and are an important context against which to cast the objective doctrine of invalidity.

7.3.2. Recent Applications of the Doctrine

In recent cases concerning the doctrine, the Constitutional Court has tried to hem the edges of objective invalidity, albeit in a way that seems difficult to square with its underlying theory. In *Department of Transport v Tasima (Pty) Ltd*[35] the court was called upon to expound on an earlier Court of Appeal case, *Oudekraal Estates v City of Cape Town*, which had laid down principles for the status of constitutionally invalid administrative acts.[36] The *Oudekraal* decision was in turn relied upon by the Constitutional Court in *MEC for Health, Eastern Cape v Kirland Investments* in supporting the view that invalid administrative decisions are binding until set aside.[37] On its face, it is not immediately obvious how this sits comfortably with objective invalidity.

---

[33] 'This Constitution is the supreme law of the Republic; law or conduct inconsistent with it is invalid, and the obligations imposed by it must be fulfilled'.

[34] 'When deciding a constitutional matter within its power, a court … must declare that any law or conduct that is inconsistent with the Constitution is invalid to the extent of its inconsistency'.

[35] *Department of Transport v Tasima (Pty) Ltd* [2016] ZACC 39, (2017) 1 BCLR 1.

[36] *Oudekraal Estates (Pty) Ltd v City of Cape Town* [2004] ZASCA 48, (2004) 3 All SA 1.

[37] *MEC for Health, Eastern Cape v Kirland Investments (Pty) Ltd* [2014] ZACC 6, (2014) 3 SA 481 [106].

The majority judgment of the court in *Tasima*, delivered by Khampepe J,[38] held the following in resolving these authorities in favour of the *Kirland Investments* view:

> [W]hen confronted with unconstitutionality, courts are bound by the Constitution to make a declaration of invalidity. No constitutional principle allows an unlawful administrative decision to 'morph into a valid act'. However, for the reasons developed through a long string of this Court's judgments, that declaration must be made by a court. … [U]ntil a court is appropriately approached and an allegedly unlawful exercise of public power is adjudicated upon, it has binding effect merely because of its factual existence.
>
> This important principle does not undermine the supremacy of the Constitution or the doctrine of objective invalidity. In the interests of certainty and the rule of law, it merely preserves the fascia of legal authority until the decision is set aside by a court: the administrative act remains legally effective, despite the fact that it may be objectively invalid.[39]

It is not exactly clear what the significance of a legal act or artefact being objectively invalid is if it still requires the imprimatur of a court to have any real consequence; similarly, it is difficult to see how the court's declaration is not having a constitutive effect here. True, it is not constitutive of the *invalidity*, but if invalidity does not necessarily entail non-bindingness[40] then it fails to simulate retrospective effect, which is one of the main points of the doctrine. It is also odd to parse the bindingness of law as an aspect of its factual rather than legal existence. Such a stance at its most extreme would seem to suggest that anything that even passingly looked like law (such as repealed provisions, perhaps?) would be binding. The factual existence of a statute inheres in the observable fact that a particular provision has been printed and publicised on the statute book. The legal existence of a statute is its ability to ground and justify legal action.

Points such as these animated the minority judgment of Jafta J,[41] who was particularly attuned to the problem of making a meaningful distinction between valid

---

[38] Joined in part by Froneman, Khampepe, Madlanga, Mhlantla and Nkabinde JJ: *Department of Transport v Tasima (Pty) Ltd* (n 35) [228]–[229].

[39] ibid [147]–[148].

[40] I discuss this point in chapter 8. The term validity is unhelpfully ambiguous. Only a certain meaning of invalidity entails non-bindingness.

[41] Also joined by Zondo J, Mogoeng CJ and Bosielo AJ: *Department of Transport v Tasima (Pty) Ltd* (n 35) [210].

and invalid laws in the context of bindingness when he proclaimed that:

> One administrative action is not capable of being invalid and valid at the same time. The notion of an invalid action remaining valid until set aside is flawed. Invalidity and validity in this context are mutually exclusive. If an invalid action were to be valid until set aside, every performance in compliance with it would be lawful. And once set aside, the invalid action would cease to exist at the level of fact. But its revocation by a court on review would not affect what had already been done in accordance with the invalid action because it would have been valid until the moment of setting it aside.[42]

Jafta J also observed that in establishing this principle in *Kirland*, some reliance was placed on recent commentary from Forsyth on English administrative law.[43] This does not of itself make the *Oudekraal*/*Kirland* approach flawed, but there is much in the difference between South African and English public law—a written constitution and strong support for hard-form judicial review not least among those differences.[44]

Whether this aspect of the judgment was correct as a matter of South African law or not, it was likely correct for the majority in *Tasima* to observe that judicial action in the form of a declaration is required before an unconstitutional law may be disregarded. This had been endorsed in several prior decisions.[45] Thus, at least as a matter of precedent, the proposition was well supported.

The Constitutional Court seemed to endorse a distinction between a law's existence/validity and its capacity to justify legal results in cases, though it somewhat unhelpfully based this point on a distinction between factual validity and legal validity.[46] The idea of factual validity is not particularly illuminating, as I argued above

---

[42] *Department of Transport v Tasima (Pty) Ltd* (n 35) [124].

[43] Christopher Forsyth, '"The Metaphysic of Nullity": Invalidity, Conceptual Reasoning and the Rule of Law' in Christopher Forsyth and Ivan Hare (eds), *The Golden Metwand and the Crooked Cord: Essays in Honour of Sir William Wade QC* (OUP 1998). Relied upon in the original *Oudekraal* case: *Oudekraal Estates (Pty) Ltd v City of Cape Town* (n 36) [29]–[30].

[44] I criticise the idea that English administrative law scholarship can be useful ported over *salva veritate* to unconstitutionality in chapter 11.

[45] *Merafong City Local Municipality v AngloGold Ashanti Ltd* [2016] ZACC 35 [42]; *MEC for Health, Eastern Cape v Kirland Investments (Pty) Ltd* (n 37) [101]–[103]; *Camps Bay Ratepayers' Association v Harrison* [2010] ZACC 19, (2011) 4 SA 42 [62]; *Camps Bay Ratepayers' Association v Harrison* [2010] ZACC 26, (2011) 4 SA 113 [85].

[46] This was the point derived from Forsyth (n 43).

in my criticism of the majority judgment. I will return to a more considered analysis of legal validity in chapter 8. For the purposes of this chapter what is important to note is that the majority's judgment in *Tasima* seems to require that where laws are found invalid, they are so found from the first time of their enactment.[47] It therefore seems that although invalidity is achieved by autonomous operation of the Constitutional supremacy clause (as it is in Canada: see chapter 4), this does not have any juridical impact until a court makes a declaration to that effect. This is the thrust of the majority judgments in both *Economic Freedom Fighters*[48] and *Tasima*.

In conclusion, South Africa holds unconstitutional legislation to be invalid and void from the point of its enactment. This is strongly evidenced by the continuing vibrancy of the doctrine of objective invalidity. However, recent judgments on that doctrine suggest that there might be increasing movement towards a more nuanced understanding of the implications of this theory. The distinction between 'factual' and 'legal' validity is in some respects an embryonic iteration of the analysis that I will carry out in chapter 8 comparing validity and applicability.

7.3.3. *Erga Omnes* Effect

There has also been some consideration in the South African courts on how the *erga omnes* effect of declarations of unconstitutionality is achieved. Although this is a more narrow point, it is still illuminating for the present discussion if the unconstitutionality applies to everyone because of precedent or because of an absence of law. Somewhat surprisingly given the different constitutional milieu of South Africa, the

---

[47] This instantiates a response to what I call the 'origin' question in Chapter 8.

[48] Where the Court had said:

> No decision grounded on the Constitution or law may be disregarded without recourse to a court of law. To do otherwise would 'amount to a licence to self-help'. … No binding and constitutionally or statutorily sourced decision may be disregarded willy-nilly. It has legal consequences and must be complied with or acted upon. To achieve the opposite outcome lawfully, an order of court would have to be obtained.

Economic Freedom Fighters v Speaker of the National Assembly [2016] ZACC 11, (2016) 3 SA 580 [74].

Court of Appeal has endorsed the same theory of precedent as expounded by Lord Diplock in *Hoffmann La-Roche*.[49] In *Kouga Municipality v Bellingan*, the Court of Appeal observed the following:

> 'Although such a decision is directly binding only as between the parties to the proceedings in which it was made, the application of the doctrine of precedent has the consequence of enabling the benefit of it to accrue to all other persons whose legal rights have been interfered with in reliance on the law which the statutory instruments purported to declare'—per Lord Diplock in *F Hoffmann-La Roche & Co AG v Secretary of State for Trade and Industry*[.] That means a declaratory order in favour of the applicants would render all prosecutions still-born …. [T]his result follows from the failure by the Council of the Municipality to pass the by-law in accordance with the empowering legislation.[50]

It is difficult to square this statement of the Supreme Court of Appeal in with either the constitutional text or the doctrine of objective invalidity. As mentioned above, section 172(1) of the South African Constitution is unambiguous with respect to declaring unconstitutional laws *invalid*. The absence of a constitutional textual basis for judicial review in England and the United States facilitates (maybe even requires) a precedent-based theory. Where there is a specific reference to invalidity, it is not at all clear why it is precedent rather than the simple non-existence of the law that is generating *erga omnes* effect. The better view would be that the court in *Bellingan* was mistaken, and the view on *erga omnes* effect therein should not be followed or viewed as correct in subsequent South African cases. The doctrine of objectivity, as it renders laws void, provides its own basis for *erga omnes* effect.

7.4. SUSPENDED DECLARATIONS OF UNCONSTITUTIONALITY

As mentioned above, both the Canadian and South African courts recognise a power to suspend declarations of unconstitutionality. However, the South African and Canadian practices are not entirely on all fours with one another. Given this difference, and because suspension in general was already explained in chapter 4, I will focus

[49] *Hoffmann-La Roche v Secretary of State for Trade* [1975] AC 295 (HL).

[50] *Kouga Municipality v Bellingan* [2011] ZASCA 222, (2012) 2 SA 95 [21]. Internal citation to: *Hoffmann-La Roche* (n 49) 365.

here on ways in which South African practice differs from that of Canada.

7.4.1. Constitutional Basis and Frequency of Suspension

One significant difference is the extent to which courts in each jurisdiction rely on the remedy. In Canada, the consensus even in the early 2000s was that suspended declarations were becoming more norm than exception in public law.[51] In South Africa, the assumption is rather that the Constitutional Court will award a declaration with immediate effect.[52] Nevertheless, Leckey observes that suspended declarations became more common after the Interim Constitution, with the Constitutional Court suspending its order in nearly thirty cases between 2000 to 2014.[53]

Under section 172(1)(b) the Constitutional Court has a broad discretion to either limit the retrospective effect of declarations of invalidity it makes, or to suspend the effect of such declarations until the legislature is given enough time to react to them. In considering a similar power under section 98(6)(a) of the Interim Constitution, the Constitutional Court reflected in the joined cases of *Bhulwana and Gwadiso* that:

> Central to a consideration of the interests of justice in a particular case is that successful litigants should obtain the relief they seek. It is only when the interests of good government outweigh the interests of the individual litigants that the court will not grant relief to successful litigants. In principle too, the litigants before the court should not be singled out for the grant of relief, but relief should be afforded to all people who are in the same situation as the litigants … [However,] we should be circumspect in exercising our powers under section 98(6)(a) so as to avoid unnecessary dislocation and uncertainty in the criminal justice process.
>
> […]
>
> As a general principle … an order of invalidity should have no effect on cases which have been finalised prior to the date of the order of invalidity.[54]

---

[51] Robert Leckey, *Bills of Rights in the Common Law* (CUP 2015) 103; Sujit Choudhry and Kent Roach, 'Putting the Past Behind Us? Prospective Judicial and Legislative Constitutional Remedies' (2003) 21 Supreme Court Law Review (2nd) 205, 228.

[52] 'The general assumption . . . is that an unconstitutional provision is invalid with immediate effect and that a party wishing the Court to suspend its order of invalidity must provide persuasive reasons for the Court to do so.' Stuart Woolman and Michael Bishop (eds), *Constitutional Law of South Africa* (2nd edn (rev), Juta 2013) 12–78.

[53] Robert Leckey, *Bills of Rights in the Common Law* (CUP 2015) 104.

[54] *S v Bhulwana; S v Gwadiso* [1995] ZACC 11, (1996) 1 SA 388 [32].

This statement reflects an early willingness of the South African courts to place restrictions on the retrospective effects of relief granted pursuant to a finding of unconstitutionality.

The Constitutional Court of South Africa strongly affirmed its willingness to use suspended declarations to their fullest effect in the *Western Cape* case.[55] Reflecting on powers in the Interim Constitution of 1993, which were of similar character to those in section 172 of the 1996 Constitution, the Court held that:

> The powers conferred on the Courts by s 98(5) and (6) [of the Interim Constitution] are necessary powers. When the Constitution came into force there were many old laws on the statute book which were inconsistent with the Constitution. If all of them were to have been struck down and all action under them declared to be invalid, there would have been a legislative vacuum and chaotic conditions. Section 98(5) and (6) enable the Court to regulate the impact of a declaration of invalidity and avoid such consequences.[56]

This most radical remedy, the suspension of the declaration of invalidity, is not an easy remedy for a litigant to get. It is generally considered appropriate only in situations where, for example, there are many possible legislative responses of sufficient complexity and nuance that the court feels that it is best left to the legislative process.[57] Alternatively, where an unworkable lacuna would be left in the area by declaring the legislation invalid, a suspension order may also be appropriate.[58]

Suspended declarations have also been considered apposite in cases where the defect in the legislation is procedural rather than substantive in nature. For example, in *Doctors for Life*, the court found that failure to engage in sufficient public consultation was unconstitutional but issued a suspended declaration to allow for the legislature to address this problem.[59]

---

[55] *Executive Council of Western Cape Legislature v President of the Republic of South Africa* [1995] ZACC 8, (1995) 4 SA 877.

[56] ibid [107].

[57] *Fraser v Children's Court Pretoria North* [1997] ZACC 1, (1997) 2 SA 218 [50]; *Dawood v Minister of Home Affairs* [2000] ZACC 8, (2000) 3 SA 936 [63]; *South African National Defence Union v Minister of Defence* [2007] ZACC 10, (2007) 5 SA 400.

[58] *J v Director General, Department of Home Affairs* [2003] ZACC 3, (2003) 5 SA 621 [21].

[59] *Doctors for Life International v Speaker of the National Assembly* [2006] ZACC 11, (2006) 6 SA 416 [114].

7.4.2. Interim Orders

An additional innovation in the South African courts is the use of the 'interim order', a remedy that was first granted in *Dawood v Minister for Home Affairs*.[60] This is an order granted during the period of suspension of the declaration to diminish the violation of rights that is technically upheld in the meantime.[61] This additional measure achieves a better balance between the competing demands of doing justice in the case at bar and deferring to legislative choice in resolving unconstitutionalities in legislative schemes. The *Dawood* case concerned the discretion of officials to refuse to grant a temporary residence permit to the spouse of a South African citizen. This was found unconstitutional.[62] The Court's order required that officials could continue to use this power but only where they considered the constitutional rights of applicants and could show compelling cause to refuse a permit.[63] The court has increasingly relied on this innovation in cases where it suspends its declaration.[64]

In some interim order cases, the Court grants orders of significant detail. Woolman and Bishop note[65] that in *Janse van Rensburg v Minister of Trade and Industry*,[66] *S v Steyn*[67] and *Zondi v MEC for Traditional and Local Government Affairs*,[68] interim orders that were of such detail that they bore resemblance to legislative provisions were awarded. Although these remedies are presented as a way for a court to avoid interference with the legislative power, it is worth noting that these interim awards are awarded in lieu of a permanent reading-in remedy; the Constitutional Court

[60] *Dawood v Minister of Home Affairs* [2000] ZACC 8, (2000) 3 SA 936.

[61] Woolman and Bishop (n 52) 9–123.

[62] A another case arose recently with near-identical facts and achieving substantially the same result: *Lawyers for Human Rights v Minister of Home Affairs* [2017] ZACC 22, (2017) 5 SA 480.

[63] *Dawood v Minister of Home Affairs* (n 60) [67].

[64] Leckey, *Bills of Rights in the Common Law* (n 53) 105.

[65] Woolman and Bishop (n 52) 9–124.

[66] *Janse van Rensburg v Minister of Trade and Industry* [2000] ZACC 18, (2001) 1 SA 29.

[67] *S v Steyn* [2000] ZACC 24, (2001) 1 SA 1146.

[68] *Zondi v MEC for Traditional and Local Government Affairs* [2004] ZACC 19, (2005) 3 SA 589.

itself has noted the mutual exclusivity of interim orders and reading-in.[69] In many respects reading in would be less of an intrusion on legislative power, as the existing words of a statute impose a greater constraint on the court. The expansive interim remedy powers must therefore be read alongside the Constitutional Court's repeated commitment to interfere to the minimum extent with the legislative power.[70]

7.5. CONCLUSION

The most distinctive remedial features of the South African practice of judicial review are its textual provision for suspended declarations of unconstitutionality, and its relatively liberal use of interim orders offering temporary solutions to difficulties generated by unconstitutional legislation during a period of suspension. These are innovative approaches to unconstitutionality and South African judicial review commends itself to comparative study because of these features.

However, against this backdrop the theory of what happens when legislation is declared unconstitutional is highly similar to other jurisdictions and thus might even be considered 'orthodox'. The doctrine of objective invalidity, and the constitutional text itself, require that legislation be deemed invalid and void *ab initio*. This reflects a similar commitment to the supremacy of the Constitution that is observable in both Ireland and Canada. The result of voidness in an 'objective' sense is also very reminiscent of the Indian and Irish practice of absolute voidness for post-constitution law.

[69] *J v Director General, Department of Home Affairs* (n 58) [21].

[70] *De Vos NO v Minister of Justice and Constitutional Development* [2015] ZACC 21, (2015) 9 BCLR 1026 [67]; *Teddy Bear Clinic for Abused Children v Minister of Justice and Constitutional Development* [2013] ZACC 35, (2014) 2 SA 168 [106]; *Dawood v Minister of Home Affairs* (n 60) [64].

Part III.

Theoretical Analysis

8 A Theoretical Framework for Unconstitutionality

CHAPTER OVERVIEW

8.1. Introduction . . . 186

8.2. Three Questions for a Theory of Unconstitutionality . . . 188

8.3. The Derivation of Unconstitutionality . . . 190

8.3.1. Judicial Intervention . . . 192

8.3.1.1. Kelsen's Argument . . . 192

8.3.1.2. Legal Realism . . . 196

8.3.2. Legislative Nullity . . . 198

8.3.2.1. Blackstone's Declaratory Theory . . . 202

8.3.2.2. 'Right Answer' Theories . . . 204

8.3.2.3. The Role of Interpretation . . . 210

8.3.3. Differentiating Legislative Nullity and Judicial Intervention 211

8.4. The Effect of Unconstitutionality . . . 215

8.4.1. Validity . . . 216

8.4.2. Applicability . . . 221

8.4.3. Invalidity does not Imply Inapplicability . . . 225

8.4.3.1. D-Validity ≠ Applicability . . . 227

8.4.3.2. N-Validity ≠ Applicability . . . 231

8.4.4. Summary and Conclusion . . . 232

8.5. The Timing of Unconstitutionality . . . 235

8.5.1. The Origin Question . . . 235

8.5.1.1. Unconstitutionality from the Future . . . 235

8.5.1.2. The Simulation of Retrospective Effect . . . 238

8.5.2. The Duration Question . . . 239

8.6. Combining Answers to the Three Questions . . . 240

8.6.1. Voidness *Ab Initio* (Ln | Va | Pa | Rt & Pr) 243
8.6.2. Applicability-Regulating Mechanisms (XX | Ap | XX | Rt & Pr) 246
8.6.2.1. Pre-Constitution Law in Ireland and India (Ln | Ap | Pa | Rt & Pr) 246
8.6.2.2. Ambica Mills in India (Ln | Ap | Pa | Rt & Pr) . . . 249
8.6.2.3. Suspended Declarations of Unconstitutionality (Ln | Ap | Pa | Rt & Pr) 249
8.6.2.4. As-Applied Challenges (Ji | Ap | Pr | Rt & Pr) . . . 250
8.6.3. Prospective-Only Mechanisms (XX | XX | XX | Pr) 251
8.6.3.1. The Legislative Nullity Variant: Ireland (Ln | Va | Pa | Pr) 251
8.6.3.2. The Judicial Intervention Variant: United States (Ji | Ap | Pr | Pr) 251
8.7. Conclusion 252

8.1. INTRODUCTION

The data generated from the study in Part II establish that unconstitutionality is a topic that can be fraught with difficulty. Challenging issues emerge in every jurisdiction. The United States has oscillated between declaring unconstitutional law to be void and declaring that there is merely an injunction against enforcing it in certain circumstances. The Indian courts have similarly struggled with this question, with the added complication of a constitutional text that is more restrictive of theoretical choice than the US Constitution. In Ireland and India, pre-constitution law is treated differently to law that was enacted under the Constitution, but in South Africa it is not. Is this merely a feature of those constitutional texts, or is there something more theoretically significant to this practice? Canada and South Africa have both developed practices of suspending declarations of unconstitutionality, but does this preserve the idea of unconstitutional law being invalid, or must it replace it with something else? Does it really make sense to call a law valid when, because of some imperfection it continues to bear, it has a sword of Damocles suspended above it that

is certain to fall? Perhaps most significantly of all: each of these jurisdictions seems to have struggled with the consequence of treating unconstitutionality as generating voidness *ab initio*.

Some of the difficulty in parsing these questions owes to imprecision in how unconstitutionality is often theorised. Consider the following quotation from a former judge of the Supreme Court of the Philippines in a textbook on constitutional law. He suggests that there are two views on declarations of unconstitutionality:

> The first is the orthodox view. Under this rule … an unconstitutional act is not a law; it confers no right; it imposes no duties; it affords no protection; it creates no office; it is, in legal contemplation, inoperative, as if it had not been passed. It is therefore stricken from the statute books and considered never to have existed at all. Not only the parties but all persons are bound by the declaration of unconstitutionality, which means that no one may thereafter invoke it nor may the courts be permitted to apply it in subsequent cases. It is, in other words, a total nullity.
>
> The second or modern view is less stringent. Under this view, the court in passing upon the question of constitutionality does not annul or repeal the statute if it finds it in conflict with the Constitution. It simply refuses to recognize it and determines the rights of the parties just as if such statute had no existence. The court may give its reasons for ignoring or disregarding the law, but the decision affects the parties only and there is no judgment against the statute. The opinion or reasons of the court may operate as a precedent for the determination of other similar cases, but it does not strike the statute from the statute books; it does not repeal, supersede, revoke, or annul the statute. The parties to the suit are concluded by the judgment, but no one else is bound.[1]

This seems attractively comprehensive and simple. It attends to differences in practice that have been observed in the jurisdiction surveys thus far. For example, it can accommodate the differences between the United States, where unconstitutional statutes have been 'revived' (the 'modern' view), and jurisdictions such as Canada, South Africa, Ireland, and India, where unconstitutionality bears strong connotations of invalidity and voidness (the 'orthodox' view). However, Cruz's analysis does not determine *why* this difference occurs. Is it because the power of judicial review in the United States is implied, rather than express? Is it because judges in the Supreme Courts of Canada and South Africa have tended to take the view that the Constitution speaks for itself, and the judge is merely a cipher for its pronounce-

[1] Isagani Cruz, *Constitutional Law* (Central Law Book Publishing 1991) 32–33.

ments? Or are all these potential explanations wrong, and is the choice simply arbitrary?

An alternative possibility is that Cruz is not even speaking about the effect of unconstitutionality. He may simply be drawing a distinction between unconstitutionality that applies with *erga omnes* effect (the 'orthodox' view) and unconstitutionality that only binds the parties to a case and other similarly situated parties, (the 'modern' view). In other words, it could be that Cruz is building his models simply as a way of differentiating the scope of unconstitutional norms and justifying this by reference to a distinction between the 'repeal' of a statute and an injunction against its enforcement. This would prioritise the scope of unconstitutional norms (the question of whom they bind, if anyone) over the juridical effects of the operation of unconstitutionality (the question of the effect of an unconstitutional norm).

In this chapter I present a more fine-grained theory of unconstitutionality. I suggest that any theory of unconstitutionality must answer three separate questions. In some cases, the answer to one question may strongly imply an answer to another. This is merely reflective of the reality that some theories are more coherent and internally consistent than others; if unconstitutionality is a matter of several interlocking parts, we should not be surprised that some parts fit better together than others. Treating the questions separately, however, promotes conceptual clarity in this area.

8.2. THREE QUESTIONS FOR A THEORY OF UNCONSTITUTIONALITY

A theory of unconstitutionality must answer at least three questions: (1) whence does unconstitutionality derive, (2) what does unconstitutionality do, and (3) when does unconstitutionality happen? Let us call these the *derivation question*, the *effects question*, and the *temporal question*. Different answers may be proffered to these questions, and so theories of unconstitutionality may vary on any of these three axes. This suggests that unconstitutionality is far from a monosemous legal idea. There is more theoretical choice open here than might first appear, and the frequently cited

view that unconstitutional law is summarily invalid and void is no sacred cow.

The derivation question may be answered in two ways: the constitution can be taken as self-regulating and self-executing, or the judge can be taken as establishing the unconstitutionality through their ruling in a case. This means unconstitutionality can be characterised as a legislative nullity or a judicial intervention. As seen in Part II, constitutional texts often create the impression that the former characterisation is to be preferred. However, this impression is not necessarily ineluctable. Unconstitutionality is often understood to be decided by this question, entailing the answers to the subsequent two questions. The three questions are certainly interrelated, but this strong entailment does not necessarily hold.

The effects question may also be answered in two ways: unconstitutional law may be invalid, or it may be inapplicable. Invalidity, strictly speaking, is a matter of system membership; it is a function of a norm's belonging to a system by dint of being created by a properly empowered authority. Applicability is a wider concept. In rough terms, it refers to the ability of a norm to cover a case, and its ability to justify a legal result in that case. Importantly, the relationship between validity and applicability is less close than might be thought. There are various legal doctrines, such as the conflict of laws, delayed commencement of legislation, and desuetude that show how they can come apart. I will argue that validity is usually sufficient and always unnecessary for applicability; that is, law can still be applicable even if it is invalid and, indeed, it may be inapplicable even if valid.[2] This implies that unconstitutional laws may be both invalid and (*pro tanto*) applicable.

The temporal question divides into two sub-questions: (1) when does the unconstitutionality start, and (2) what periods does the unconstitutionality cover? These are questions of temporal origin and temporal duration. The question of origin can be resolved either by having the declaration of unconstitutionality take immediate

[2] It is important to clarify that this is a claim about the logically necessary relationship between validity and applicability. In many cases, invalidity will *contingently*, but not necessarily, imply inapplicability. The crucial point in this distinction is that if a connection is contingent then it is not ineluctable and, as a consequence of this, the effects of invalidity may be less severe.

effect, or by characterising the unconstitutionality as having its origin at some point in the past (suggesting that the constitutional norm grounding the unconstitutionality pre-existed the case). The duration question, finally, can be answered either by having unconstitutionality take effect both prospectively and retrospectively, or prospectively only.

8.3. THE DERIVATION OF UNCONSTITUTIONALITY

> *Whence does unconstitutionality derive? In virtue of what power, or failure of the exercise of power, is unconstitutionality established?*

The derivation question shares much in common with a burgeoning philosophical interest in an enquiry into a particular kind of dependence relation that is often captured by the 'in virtue of' relationship. An example of this type of relationship is the statement: '*in virtue of* the 58th American presidential election held on the 8 November 2016, Donald Trump became the 45th President of America'. This type of dependence between things is often called 'grounding'.[3] Some legal philosophers have also taken an interest in this analysis.[4] However, the term 'grounds' is often used in legal parlance to describe a specific *instance* of an 'in virtue of' relationship. It might be said that article X of the Constitution is the ground of the unconstitutionality of statutory provision Y. This is what would often be understood by the 'ground of unconstitutionality'. I use the term 'derivation' to describe the general grounding relationship that obtains between unconstitutionality *as a concept*, and whatever that concept depends on.

However, to describe this claim as conceptual might overstate its breadth. What is of interest here is not the broad philosophical claim 'in virtue of what does law ex-

[3] General overviews of the literature include: Kelly Trogdon, 'An Introduction to Grounding' in Miguel Hoeltje, Benjamin Schnieder and Alex Steinberg (eds), *Varieties of Dependence: Ontological Dependence, Grounding, Supervenience, Response-Dependence (Basic Philosophical Concepts)* (Philosophia Verlag 2013); Michael Clark and David Liggins, 'Recent Work on Grounding' 72 Analysis 812.

[4] Triantafyllos Gkouvas, 'Resisting Perspectivalism About Law: The Scope of Jurisprudential Disagreement' (2017) 8 Jurisprudence 205; Jules Coleman, 'The Architecture of Jurisprudence' (2011) 121 Yale L J 2.

ist?' The answer to this might be social facts, moral facts, or something in between.[5] I seek to answer a much more limited and particular conceptual question: 'in virtue of what is law unconstitutional?' There are broadly two answers to this categorical derivation question: (1) judicial intervention, and (2) legislative nullity.

The judicial intervention answer maintains that judges 'make' things unconstitutional; they establish unconstitutionality through the speech-act of giving judgment in the case. The legislative nullity answer maintains that unconstitutionality has nothing to do with judges, at least constitutively; rather, unconstitutionality is a failure of the legislature to legislate properly, and this failure is constituted by automatic operation of the constitution. This latter characterisation can be pragmatically useful for judges useful insofar as it effaces judicial agency; it suggests that the Constitution itself somehow acts as the agent establishing the unconstitutionality. Almost a century ago, Dewey neatly encapsulated the logical distinction between these enterprises:

> The logic of exposition is different from that of search and inquiry. In the latter, the situation as it exists is more or less doubtful, indeterminate, and problematic with respect to what it signifies. It unfolds itself gradually and is susceptible of dramatic surprise; at all events it has, for the time being, two sides. Exposition implies that a definitive solution is reached, that the situation is now determinate with respect to its legal implication. Its purpose is to set forth grounds for the decision reached so that it will not appear as an arbitrary dictum, and so that it will indicate a rule for dealing with similar cases in the future.[6]

Importantly, for the purposes of this question, what 'unconstitutionality' itself does must be kept ambiguous. It could refer to various modifications to legal norms that will be discussed further below under the 'effects' question. What is important for the derivation question is not what unconstitutionality *does* to legal norms; rather, it is the question of *how* the operation of unconstitutionality is performed on legal norms.

---

[5] Mark Greenberg, 'The Moral Impact Theory of Law' (2014) 123 Yale L J 1289; Gideon Rosen, 'Metaphysical Dependence: Grounding and Reduction' in Bob Hale and Aviv Hoffmann (eds), *Modality: Metaphysics, Logic, and Epistemology* (OUP 2010) 113–14; Mark Greenberg, 'How Facts Make Law I' in Scott Hershovitz (ed), *Exploring Law's Empire: The Jurisprudence of Ronald Dworkin* (OUP 2008) 251.

[6] John Dewey, 'Logical Method and Law' (1924) 10 Cornell L Rev 17, 24.

8.3.1. Judicial Intervention

Let us take the judicial intervention answer to the effects question first. What is important for this view is that legislation that is constitutionally defective *still counts as legislation* until the court exercises its review power. Thus, this answer to the derivation question must establish some convincing means of counting defective exercises of the legislative power as exercises of the legislative power nevertheless. Difficulties on this score were already seen in chapter 6 in the context of the disagreement in the Supreme Court of India as to whether some instance of unconstitutionality went to legislative competence and others did not. As shall be seen, the view that the legislative power can be validly exercised in a manner that contradicts the constitution is difficult to establish by theoretical argument.

8.3.1.1. Kelsen's Argument

Kelsen has suggested that an exercise of the legislative power resulting in the production of a constitutionally invalid norm is nevertheless still a valid exercise of the legislative power under the terms of a 'tacit alternative clause' in the Constitution. This alternative clause is not the express legislative power, but rather an implicit constitutional power that operates as a kind of 'catch-all' where legislative procedure is not followed. In effect, the tacit alternative clause is a non-textual, default legislative power. This tacit alternative clause is then subject to a special repeal[7] mechanism by the courts, which is activated when a statute is found to be unconstitutional.

To see how and why Kelsen arrives at this conclusion, we must first notice that there is something of a tension in Kelsen's thought between the following two propositions, both of which he seems to want to support:[8] *(i)* a statute that falls outside the scope of its parent norm is invalid, and therefore legally non-existent;[9] and

[7] It is notable that Kelsen seems to think that unconstitutional norms are repealed. I argue below that this is a tempting, but misleading, view to take.

[8] This argument is more fully cashed out in Pablo Navarro and Jorge Rodríguez, *Deontic Logic and Legal Systems* (Cambridge University Press 2014) 118–22.

[9] Hans Kelsen, *Pure Theory of Law* (Max Knight tr, University of California Press 1967) 271.

(ii) certain legal organs within the legal system can evaluate the compatibility of lower-order norms with their parent norms and, *unless an empowered organ makes a decision to the contrary*, these lower-order norms are binding.[10]

On the face of it, Kelsen cannot accept that an unconstitutional law is a valid norm, per *(i)*. However, he seems to also accept that a norm can continue to produce normative effects until it is evaluated by a properly empowered organ as being inconsistent with a higher norm, per *(ii)*. So, Kelsen seems to be caught in a bind between not recognising the unconstitutional law as a norm at all, and having to recognise it as a norm until it is evaluated and stricken by an appropriately empowered organ (in this case, a court).

In the general run of things, Kelsen does not seem to be willing to accept voidness *ab initio* as possible.[11] A notable exception to this is where a legislative power is exercised by a body other than the designated legislature—say, a private citizen or group enacts a 'statute'. In this case, Kelsen is prepared to declare the purported statute null and void:

> It is undeniable that there are cases in which something, especially a command that claims to be a legal norm, need not be so regarded by anybody without the legal order authorizing everybody to maintain such a position, that is, without, in fact, an act of nullification rendered by a special organ being necessary—for example, if a patient in an insane asylum issues a 'statute'. If we assume that in these cases nullity exists *a priori*, then such nullity falls outside the sphere of law.[12]

Importantly, Kelsen does not think that this result is capable of being arrived at 'legally': the nullity here is not one which is determined by legal rules. But if the nullity in question is not a legal nullity, then what is it? It would seem wrong to say that it is a *factual* nullity that Kelsen's insane asylum patient has issued a 'statute'. Kelsen thus has difficulty drawing a tenable distinction between total authority failures, such as an individual merely stipulating that they have legislative power, and borderline authority failures such as a correctly constituted parliament following

[10] Hans Kelsen, 'Judicial Review of Legislation' (1942) 4 Journal of Politics 183, 189–91.

[11] ibid 190.

[12] Kelsen, *PTL* (n 9) 277–78.

constitutional procedure but enacting an instrument that offends against a constitutionally protected right. The idea that Kelsen wants to express here is that some (total) failures of authority are void, whereas other (partial) failures of authority are merely voidable.

Kelsen made the point that judgments declaring voidness (or, more accurately for Kelsen, voidability) are constitutive and not declaratory in a comparative study between the US and Austrian constitutions. Of the US practice of voidness *ab initio* he said:

> It is especially impossible to consider a statute enacted by the constitutional legislator as absolutely null or 'void *ab initio*.' Only courts have the power to decide the question whether a statute is unconstitutional. If another person refuses to obey a statute enacted by the constitutional legislator because he believes the statute to be unconstitutional he acts on the risk that the competent court considers his conduct as illegal because the court regards the statute as constitutional. From a legal point of view only the opinion of the court is decisive. Therefore the statute must be considered valid so long as it is not declared unconstitutional by the competent court. Such a declaration has, therefore, always a constitutive and not a declaratory character.[13]

Although at time of writing Kelsen may have been correct to suggest that the United States practised voidness *ab initio*, this is likely no longer the case.[14] Nevertheless, Kelsen's opposition to voidness *ab initio* is clear. How, then, does Kelsen propose to rationalise this practice and preserve the purity of his *Pure Theory*? In short, as I indicated earlier, Kelsen makes the surprising claim that unconstitutional statutes are in fact *valid*.[15] Recall that Kelsen posits two ways a statute can constitutionally come into being: a direct method, and a tacit 'alternative' method that is

---

[13] Kelsen, 'Judicial Review of Legislation' (n 10) 190.

[14] However, he earlier makes a point that is much closer to the modern practice, though his argument in general is not sufficiently fine-grained in its distinction between validity and applicability to carry this much further and it is inattentive to the incompatibility between this approach and a pure void *ab initio* approach:

> The fact that a law-applying organ declares a general rule unconstitutional and does not apply it in a given case means that this organ is authorized to invalidate the general rule for the concrete case; but only for the concrete case, since the general rule as such—the statute, the ordinance—remains valid and can, therefore, be applied in other concrete cases.

ibid 185.

[15] Kelsen, *PTL* (n 9) 271–76.

taken to be used when the direct method is not followed. Statutes created by the alternative method are then subject to a special repeal mechanism (ie, a competent court striking them down). Thus, there can never be a failure of legislation and a keystone of the legislative nullity model, the idea that a valid legal norm was never created at all, fails to obtain.

This resolution comes at a high price to other parts of Kelsen's theory, however.[16] For instance, it does not seem to support the view of the legal system as a hierarchy with the constitution at the apex. If the constitution can authorise any and all norms below it, as they must be authorised by one mode of legislation or another, which are jointly exhaustive, then it no longer poses any meaningful kind of restriction on enactment. A lack of such restriction renders the idea of a hierarchy pointless: if superordinate norms cannot restrict subordinate ones, how do they exist in any meaningful hierarchy? This begins to unravel the systematicity of law, which seems to be something Kelsen is at pains to assert.

There is also a problem of normativity.[17] Kelsen's tacit alternative clause seems to suggest that legislators have absolute discretion. If the norms enacted by legislatures are always valid under the constitution, then the constitutional provision providing for the power to legislate has the following content: 'ϕ [or $\neg\phi$]'[18] (the tacit nature of the alternative clause represented by square brackets). This formulation cannot be properly normative. Legal norms must necessarily exclude some class of conduct (for example, conduct that would contradict a legal obligation) to function adequately and provide genuine guidance.[19] However, a norm of the content 'ϕ [or $\neg\phi$]' cannot exclude any conduct. A solution to unconstitutional laws that dissolves the normativity of law in this way is quite plausibly a cure worse than the disease. Thus, I suggest Kelsen's rejection of voidness *ab initio* fails.

---

[16] See generally: Navarro and Rodríguez (n 8).

[17] ibid 121.

[18] 'Phi or not-Phi'. Phi is here, and elsewhere in this thesis, used as a representation of a generic and nondescript action.

[19] HLA Hart, *The Concept of Law* (2nd edn, OUP 1994) 6.

8.3.1.2. Legal Realism

A plausible alternative to this flawed Kelsenian strategy would be a kind of legal realist view. To my knowledge, no bespoke legal realist has made this specific argument in constitutional theory. It is not difficult to come up with an argument in support of the model that would be attractive to such a theorist, however. The argument is perhaps among the most simple for the judicial intervention view:

Realist Argument for Judicial Intervention

Since we are making this argument from the realist's assumed point of view, we can take as given that judges have some creative role in shaping the law and are not just formalist reasoners. One of the ways in which a judge might make law is by derogating existing law; this is because to derogate a law, the judge need only issue a new legal norm that is the negative of the derogated norm. Therefore, at least some norms will cease to exist when a judge says so.

This may strike the reader as an obvious, even banal, position. It will be even more surprising, then, to discover that few legal systems have adhered to it. A few philosophers of law have even advanced views on voidness and invalidity that would notionally oppose it, at least insofar as it might apply to disputes about legal validity.[20] It is perhaps because of the politically contentious nature of admitting to judicial law-making that this position has not been taken up more readily in the common law world. In constitutional adjudication, the political stakes are often notably higher than they are in other contexts of legal adjudication. When a court undertakes judicial review of legislation, it is not just a judge interfering with judge-made law; rather, it is a judge altering policy decisions of the legislature that have been recorded in statute. This may stretch the limits of the creativity a legal realist may be willing to permit.

[20] 'We normally do not have to wait for a court decision; we usually know that a rule will be valid if it satisfies certain conditions.' Stephen Munzer, 'Validity and Legal Conflicts' (1973) 82 Yale L J 1140, 1172–73.

Another argument against this simple realist view is that it is descriptively inaccurate in the sense that it fails to capture how some constitutional orders conceptualise what is happening when a statute is declared unconstitutional: the officials in the order do not see themselves as engaging a creative function.[21] Of course, descriptive inaccuracies are not an issue if a theory's aims are purely normative in suggesting a *better* way to do things. This does not inoculate the realist view from other criticisms, such as the challenge that it can be capricious with regard to who can avail of the benefit of a holding of unconstitutionality and who cannot.

This realist, to be consistent with their view that laws only become unconstitutional with the judge's say-so, will be compelled to favour prospectivity along the lines of the *Linkletter* case law. Again, as was seen in chapter 3, this introduces unwelcome elements of caprice: why ought constitutionality depend on the ability of citizens to bring forward issues to judges, and how far does the benefit of that unconstitutionality extend once apprehended? Hinging unconstitutionality solely on judicial declaration makes it so that the state may get away with conducting itself in a way that would or should be unconstitutional if it ever came before a judge, but cannot be so classified *until* it arises in that way. This presents a serious problem for the individual litigant. Consider a prisoner sentenced to death under a constitutionally vulnerable law. They may get advice from their lawyer that they would have a good chance of success were they to take a challenge to the problematic law, but they would also have to be advised that it will not matter much, as they will be

[21] See the criticism of the descriptive accuracy of exclusive legal positivism, which similarly conceives of constitutional norms as *empowering judges* to change the law, in: Matthew Grellette, 'Legal Positivism and the Separation of Existence and Validity' (2010) 23 Ratio Juris 22, 29–30.
See also the various degrees of adherence to Blackstone advocated in the jurisdictions under study. Those that favour legislative nullity tend to leave some scope for the declaratory theory: *A v Governor of Arbour Hill Prison* [2006] IESC 45, [2006] 4 IR 88 [35]; but see the analysis in Robert Noonan, '*Stare Decisis*, Overruling, and Judicial Law-Making: The Paradox of the *JC* Case' (2017) 57 Ir Jur 119; *Canada (Attorney General) v Hislop* 2007 SCC 10, [2007] 1 SCR 429; *R v Jordan* 2016 SCC 27 (CanLII), [2016] 1 SCR 361 [93].
Those that are not so concerned with constitutional supremacy tend to reject the Blackstone theory in less equivocal terms: *Golaknath v State of Punjab* [1967] INSC 45, [1967] 2 SCR 762. South Africa is an outlier here as a jurisdiction that consistently asserts constitutional supremacy, but which has also rejected Blackstone: *Fourie v Minister of Home Affairs* [2004] ZASCA 132, (2005) 1 All SA 273 [23].

sentenced to death anyway even if they win their case. It would be a very altruistic prisoner indeed that occupied their final months/years in the courts pursuing such a challenge for the greater good of the community going forward.[22]

Thus, although the legal realist model for a judicial intervention approach to unconstitutionality has better prospects than the Kelsenian argument, it faces some significant practical and institutional concerns. I return to these in greater detail in chapter 10.

8.3.2. Legislative Nullity

The alternative answer to the derivation question is to characterise unconstitutionality as a legislative failure.[23] The basic effect of the legislative nullity answer is that the impugned law never really existed at all; it is held to have been 'void *ab initio*'. This seems to entail the corollary effect that actions taken on foot of the void law are not legal actions. The remedy for a successful claim, under this theory, is thus not localised to the successful claimant as an individual. An unconstitutional statute is taken to be flawed in ways that are general and not plaintiff-specific, and thus the consequences of invalidity are given *erga omnes* effect. This conclusion is normally reached by some application of the following argument, or a cognate argument:[24]

Legislative Nullity

Assume that the legislative power is created by the constitution. This will be done subject to certain procedural and substantive constraints; legislation

[22] A case not entirely unlike this hypothetical arose recently in Ireland: *AB v The Clinical Director of St Loman's Hospital* [2018] IECA 123.

[23] The phenomenon of voidness *ab initio* is sometimes attributed to the declaratory theory of law: Brian Fitzgerald, 'When Should Unconstitutionality Mean "Void Ab Initio"' (1994) 1 Canberra L Rev 205, 206; Ciarán Lawlor, 'Troubling Times: Intertemporal Law and Theories of Approach to the Effects of Unconstitutionality' (2008) 7 COLR 71, 76.

[24] How descriptively accurate this argument is will, of course, vary from legal system to legal system. I do not mean to present this as a necessary truth about law; rather, I am simply expounding on what I think is a reasonably common pattern of argument that underscores thinking around unconstitutionality in at least *some* jurisdictions.

> must have a certain form, be produced through a certain process, etc. The constitution will have a term, implicit or explicit, that exercises of constitutional powers must be consistent with the whole of the constitution, not just the provisions that establish those powers. It therefore follows that an exercise of the legislative power cannot violate other provisions of the constitution. Tautologically, only the legislative power may create law through legislation. It must therefore be the case that any purported exercise of the legislative power violating a term of the constitution is not an exercise of the legislative power. It thus creates no law and this failed exercise is legally void *ab initio*.

Given the terms in which this argument is phrased, it is most suited to jurisdictions that have a written Constitution—particularly one in which a bill of rights is enshrined—as this establishes clear limitations of the legislative power based on rights concerns. Such jurisdictions will often have an explicit separation of powers with certain provisions providing limits on the legislative power. It is important also to note from the outset that another assumed corollary of the legislative nullity theory is that invalid legislation is legally non-existent and inapt to justify legal decisions. It is, however, possible to challenge this assumption. I take up this task below in answering the question of the effect of unconstitutional law. It is theoretically possible to construct a version of the legislative nullity view that goes to applicability rather than validity; this said, such a reconstruction would fight an uphill battle against the Constitutional text in many cases. Validity is a matter of a norm having a proper source within a legal system, and it may be difficult to find such a source if one grants the legislative nullity view its assumption that the processes and fetters on the legislative power must be strictly constructed. I return to these issues later; for present purposes, I follow the orthodox formulation of this view, which is that 'legislation' lacking legislative imprimatur is legally invalid, null, and void.

An interesting feature of this theory is the assertiveness it requires of the judiciary in policing exercises of the legislative power. There is a significant difference to note, however, between what judges are doing here, which I will call 'negation',

for lack of a better term, and repeal of the legislation by a subsequent legislative act. As was seen in some previous chapters, judges in some jurisdictions have occasionally conflated these concepts.[25] There are two possible ways of derogating a norm, and repeal and negation are the terms I use to describe these different mechanisms:

> *Repeal:* repeal of a legal norm is accomplished by the legislature issuing a new legal norm that either itself contradicts and supersedes the old legal norm or issues a command not to apply the derogated norm.
>
> *Negation:* the negation of a legal norm is accomplished by the judiciary finding the norm defective in a way that makes it legally insignificant. The norm may continue to exist as a norm, but it never had any existence as a *legal* norm.

A significant difference between these methods of derogation is that repeal grants that the derogated legal norm had some prior significant existence as a law. After all, one cannot repeal what does not exist.[26] The comparison between the legislative nullity view and repeal is inauspicious in other ways as well. For example, legislative repeal of provisions is usually limited to having prospective effect only. As we

---

[25] A particularly stark example was the Indian courts' experience in *Keshavan Madhava Menon v State of Bombay* [1951] INSC 3, [1951] SCR 228.

[26] This observation has been made by Navarro and Rodríguez:

> When a given norm is authoritatively declared to be unconstitutional, it is undeniable that it ceases to produce certain legal effects. But this does not necessarily mean that invalid norms do not 'exist.' In fact, *if they did not exist, there would be no need for specific organs to repeal them*. It is precisely because they have some sort of existence that they pose a serious problem for legal theory.

Navarro and Rodríguez (n 8) 119–20. Emphasis added.
If, as seems most sensible within the context of the rest of their writing and thought, Navarro and Rodríguez are here implying that it is necessary for the *legislative* organ to repeal unconstitutional laws then their claim seems true. This is more than likely the intended reading, as they also think that judges regulate applicability, not validity, in claims of unconstitutionality. It seems reasonable to extrapolate from this that Navarro and Rodríguez may view the legislature as the only legal institution that can modify the validity of laws. An alternative, though less likely, interpretation would be that the authors could be taken to claim that invalid norms, which are not *made* invalid by a court decision but are simply *declared* invalid by that decision, require the courts specifically to repeal them. This would be an unusual characterisation of what is going on when a court declares a norm unconstitutional for the reasons I supplied above in my critique of Kelsen's theory.

shall see when considering the temporal questions below, this is something that unconstitutionality, whether in its judicial intervention or legislative nullity guise, is technically agnostic on; however, as instantiated in the jurisdictions studied in this thesis, the rule is generally that unconstitutionality takes both retrospective and prospective effect.

An important commonality to note between the answers to the derivation question is that *neither* of them entail that the judiciary repeals an unconstitutional statute. In the legislative nullity case, it is simply that the legislation failed to exist *qua* legislation. There is thus nothing to repeal. As Hart memorably put the point: 'nothing which legislators do makes law unless they comply with fundamentally accepted rules specifying the essential lawmaking procedures'.[27] In the judicial intervention case, it is rather that the legislation is disapplied in certain, potentially very sweeping, circumstances, but it has every chance of finding reapplication should those circumstances change (eg, the decision declaring the unconstitutionality is overruled) or in different circumstances (eg, those that would not be covered under ordinary application of the rules of *stare decisis* to subsequent cases). Importantly, this latter claim is technically contingent on the power of the judiciary being one determining *applicability*. This presumes an answer to the 'effects' question below. Where the judicial intervention view has been taken in the jurisdictions surveyed in this study, it has been accompanied by this answer. This is not necessarily invariable; Constitutional Courts in civil law jurisdictions occasionally have powers of repeal that are effectively judicial intervention combined with a power to alter the validity (membership of the legal system) of certain legal norms. However, even if the judicial intervention view was combined with a validity analysis, it would still likely differ from repeal in terms of retrospective effect.

---

[27] HLA Hart, 'Positivism and the Separation of Law and Morals' in HLA Hart, *Essays in Jurisprudence and Philosophy* (Clarendon Press 1983) 59.

8.3.2.1. Blackstone's Declaratory Theory

One historically popular view supporting a judicial intervention approach to the derivation question is the declaratory theory of precedent commonly attributed to Blackstone. This view maintains that judges do not create law in their adjudications; they merely apply pre-existing law.[28] This way of thinking about the judicial role is deeply ingrained in reflections on precedent in the common law; indeed, it has been observed that precedent is 'what remains of the pre-realist vision of "fundamental principles of law; which, though legislators may depart from, yet judges are bound to observe".'[29] There is thus an intricate connection between the legal tradition of adherence to precedent and the Blackstonian view on the role of the judge.

Although the declaratory theory is widely discredited in modern writing,[30] it has an unusual resilience insofar as judges are often reluctant to acknowledge a law-making function that more expansive views of precedent would afford them. Though judges often facially discredit the model in a comparable manner to academics, there is still some desire to maintain it.[31] This is likely due to both its tendency to obfuscate political or policy-oriented decision-making in jurisdictions that historically have not given that role to the judiciary, and the ease with which it achieves retrospective effect in judicial decisions, which might be independently morally valuable. There is an element of having one's cake and eating it too in this: judges simultaneously discredit the model as unrealistic but wish to retain it only insofar as it is expedient to do so for explaining, or obviating the need to explain, the delicate politics of the judicial role in modifying the law.

---

[28] William Blackstone, *Commentaries on the Laws of England* (vol 1, Cavendish 1766) 69.

[29] Laurence Claus, 'Montesquieu's Mistakes and the True Meaning of Separation' (2005) 25 OJLS 419, 434. Internal citation to: Blackstone (n 28) ch 7, 259.

[30] Lord Reid, 'The Judge as Lawmaker' (1972–1973) 12 J Soc Public Teachers 22; Richard McManus, 'Predicting the Past: The Declaratory Theory of the Common Law –From Fairytale to Nightmare' (2007) 12 Judicial Review 228; Peter Mirfield, 'A Challenge to the Declaratory Theory of Law' (2008) 124 LQR 190.

[31] *In Re Spectrum Plus Ltd (In Liquidation)* [2005] UHKL 41, [2005] 2 AC 680 [34]; *Canada (Attorney General) v Hislop* 2007 SCC 10, [2007] 1 SCR 429.

The following is a condensed argument for the declaratory theory of precedent:

Declaratory Theory of Precedent

Assume that judges cannot create law; the role of the judge is merely to apply the correct law. This implies a concomitant power to disapply incorrect law. If this is true, then two further theses follow: *(i)* legal propositions (or at least the truthmakers for those propositions) exist independently of, and anterior to, the judgment that declares the law, and *(ii)* judges may be mistaken as to the law, as determined by these truthmakers.

It follows from the above that a statement that does not express a true legal proposition is not a legally binding precedent and, as such, it may be overruled. Since it was not judicial pronouncement that constituted the truth of the legal proposition, the precedent was wrong for the entire duration of the true proposition's existence. This is assumed to be time immemorial. This being the case, the legally false statement sets no precedent in law and was wrong from its inception. This view is often expressed through judicial tests for overruling that concentrate on the 'wrongness' or 'erroneousness' of prior precedent.

Where a judge is adjudicating on a constitutional issue, the Blackstonian theory requires that the constitution has some objective meaning and this meaning is what the judge seeks to understand and apply.[32] Applying this meaning in a case concerning the unconstitutionality of legislation, the judge renders clear that the statute is unconstitutional. However, as under this theory judges have the power neither to create nor destroy a law, it follows that something *else* must be performing the action of unconstitutionality. This would be the objective meaning of the Constitution itself. In other words, this Blackstonian view, in a constitutional context, requires that the Constitution regulates its own meaning, judges discover what

[32] In this respect, it is not dissimilar to the 'Fixation Thesis' espoused by Solum as one of the characteristic claims of American originalism: Lawrence Solum, 'The Fixation Thesis: The Role of Historical Fact in Original Meaning' (2015) 91 Notre Dame L Rev 1.

that meaning is, and then must announce the consequences for any laws that contradict this meaning. The Constitution regulating unconstitutionality on its own terms is a particularly strong theme in Canada and South Africa.

8.3.2.2. 'Right Answer' Theories

A relatively more current instantiation of Blackstonian thinking is to be found in what might be called 'right answer' theories of law. These theories claim that there are unique, true answers to legal questions. There are two particularly significant attempts in modern jurisprudence arguing for a theory of law that incorporates this feature: Dworkin's interpretivism, and Moore's metaphysical legal realism. In this chapter, I will focus primarily on Dworkin's work as the original 'right answer' thesis. I will return to Moore's thesis, and some more general philosophical issues around right answer theses and objectivity, in chapter 10.

In Dworkin's own words, to hold the 'right answer' thesis is to maintain that 'in most hard cases there are right answers to be hunted by reason and imagination'.[33] Like Blackstone, Dworkin claims that legal disputes will have pre-existing and discoverable right answers. In Dworkin's case, these answers are derived from interpreting principles in the law to construct a theory that best fits and justifies those rules in the existing skein of the law. This is then applied to the case at bar by the judge, thus avoiding 'retrospective' application of law, as the 'new' law is held to have existed at the time of the wrong. Given this effect, it is unsurprising that some have characterised Dworkin's view as 'neo-Blackstonian'.[34]

As a preliminary matter, it should be noted that one of the difficulties in dealing with any aspect of Dworkin's philosophy is determining what Dworkin has canonically said on a topic.[35] This difficulty has also been noted in the particular context

---

[33] Ronald Dworkin, *Law's Empire* (Hart Publishing 1998) viii–ix.

[34] Kermit Roosevelt, 'A Little Theory is a Dangerous Thing: The Myth of Adjudicative Retroactivity' (1999) 31 Conn L Rev 1075, 1104–06.

[35] As Penner has put it:

of the 'right answer' thesis.[36] In dealing with Dworkin's views for the purposes of this thesis, I will focus primarily on his earliest work developing the right answer thesis: 'Hard Cases'.[37] Taking Dworkin's earliest work has the advantage of being relatively unmoored from Dworkin's more general theory of law, which was gradually developed over a number of subsequent works. Hard Cases is intended to stand alone as a challenge to the positivist model of adjudication rather than an assertion of Dworkin's own legal theory.[38] This gives it greater generality; it would be possible to accept parts of Hard Cases even if one were not a committed Dworkinian.

In Hard Cases, Dworkin draws a distinction between decisions made on grounds of policy, and those made on grounds of principle. Arguments of policy, Dworkin says, will 'justify a political decision by showing that the decision *advances or protects some collective goal of the community as a whole*'.[39] Arguments of principle, then, will 'justify a political decision by showing that the decision *respects or secures some individual or group right*'.[40] There are two obvious differences between policy and principle here: policy relates to 'goals' and spans over entire communities; principle relates to 'rights' and extends only to individuals or small collectives. Goals are matters of aspiration and desire, whereas rights are matters of fixed legal enti-

---

> [I]t is not always easy to say with precision exactly what Dworkin believes or intends about certain issues given his propensity to revisit these issues many times in his large body of work. He has framed his views in many different ways, and it is not always easy to say with assurance that Dworkin's actual view is this rather than that.
>
> James Penner, 'Law and Adjudication: Dworkin's Critique of Positivism' in James Penner, David Schiff and Richard Nobles (eds), *Jurisprudence & Legal Theory: Commentary and Materials* (OUP 2002) 334

[36] 'For a long time, the idea most closely associated with Dworkin's work in legal theory was the "right answer thesis", the claim that all (or almost all) legal decisions have a unique right answer. It is interesting to note some of the ways that the presentation of this view, and attacks on it, have changed over time.' Brian Bix, *Jurisprudence: Theory and Context* (6th edn, Sweet & Maxwell 2012) 98.

[37] Ronald Dworkin, 'Hard Cases' (1975) 88 Harv L Rev 1057. This view has been characterised as a 'strong' right answer thesis: Joshua Geller, 'Truth, Objectivity, and Dworkin's Right Answer Thesis' [1999] UCL Jurisprudence Rev 83.

[38] For an overview of this debate, see: Anne De Moor, 'Nothing Else to Think? On Meaning, Truth, and Objectivity in Law' (1998) 18 OJLS 345.

[39] Dworkin, 'Hard Cases' (n 37) 1059. Emphasis added.

[40] ibid 1059. Emphasis added.

tlement. Policy arguments, then, are chiefly the domain of the legislature whereas arguments of principle are how courts resolve disputes.

Dworkin further suggests that not only should judges be restricted to arguments of principle, but also that they should be unoriginal in their adjudication; that is, their adjudications should not create new legal rights. Two motivating factors support this assertion. (1) as they are not elected, judges are unaccountable in a way legislators are not and so should not have a role in law-making, and (2) judicial creativity entails that a losing party to a case is punished retroactively, according to a duty created after the event, rather than a duty that existed at the time of the event giving rise to the cause of action.[41]

These priors established, Dworkin then moves to advance the 'rights thesis', which is that 'judicial decisions enforce existing political rights'.[42] This has an important effect on how precedent is construed:

> If the [rights] thesis holds, then institutional history acts not as a constraint on the political judgment of judges but as an ingredient of that judgment, because institutional history is part of the background that any plausible judgment about the rights of an individual must accommodate. Political rights are creatures of both history and morality: what an individual is entitled to have, in civil society, depends upon both the practice and the justice of its political institutions.[43]

The political rights that the rights thesis seeks to enforce are thus a complex cocktail of moral requirements and institutional history. This suggests a universality or generality that is somewhat ill at ease with Dworkin's later qualification that he believes that the rights thesis holds symmetrically only in civil cases but asymmetrically in others.[44] In a criminal case, Dworkin suggests that the accused has a right to a decision in his favour if he is innocent, but the state has no right to a conviction if he is guilty. Dworkin seems to think that this stems from an assumption

[41] Dworkin, 'Hard Cases' (n 37) 1061.

[42] ibid 1063.

[43] ibid 1063.

[44] ibid 1077.

that only in civil cases is there a presumption that one party has a right to win.[45]

There is one more feature of Dworkin's theory to cover, and that is the role played by the epistemically ideal judge: Hercules. Dworkin deploys Hercules as a judge that will be drawn to the rights thesis but, importantly, he is a judge who is ideal at applying that thesis as well. Dworkin thinks that hard cases pose political questions, and such questions will require that the adjudicator is able to reason not just about individual rules, but also about the enterprise in which they are engaged more broadly.[46] Hercules is a judge possessed of exceptional skill that enables him to do just this. Dworkin shows us how Hercules will reason through hard cases that implicate the constitution, statutory law, and the common law.[47] In each case, Hercules will be required to balance fitting the law within the scheme of legal propositions that have come before, as well as viewing the question in wider political context. Some cases will require him to theorise about what the legislature intended in pursuing a particular action (legislation), others may not require this (the common law). In all cases he must construct an abstract general theory of the constitution, legislative scheme, or common law rule that is capable of fitting and justifying all the legal rules in that area with one another.[48]

Two notable features, or limitations depending on perspective, of Hercules' adjudication are worth noting. First, Hercules' own prior precedents will be influential in his decision-making on the axis of fit.[49] Second, there seems to be a possibility of genuine disagreement between Hercules and other judges on some theoretical issues:

> Suppose Hercules … proposed to [construct his legal theory] in advance …. He would

[45] ibid 1077.

[46] ibid 1082.

[47] ibid 1083–1101.

[48] 'You will now see why I called our judge Hercules. He must construct a scheme of abstract and concrete principles that provides a coherent justification for all common law precedents and, so far as these are to be justified on principle, constitutional and statutory provisions as well.' ibid 1094.

[49] ibid 1095.

> begin … by setting out and refining the constitutional theory he has already used. *That constitutional theory would be more or less different from the theory that a different judge would develop, because a constitutional theory requires judgments about complex issues of institutional fit, as well as judgments about political and moral philosophy, and Hercules' judgments will inevitably differ from those other judges would make.*[50]

The italicised text seems to suggest that there are certain aspects of constitutional theory construction—an aspect of a general theory of a legal system for the judge—that come down to subjective judgment. Different judges may arrive at different theoretical conclusions and, it seems, we do not need to give Hercules any epistemic priority over these matters. But this seems difficult to square with a later statement in which Dworkin holds that:

> It is perfectly true that in some cases Hercules' decision about the content of [] community morality, and thus his decision about legal rights, will be controversial. This will be so whenever institutional history must be justified by appeal to some contested political concept, like fairness or liberty or equality, but it is not sufficiently detailed so that it can be justified by only one among different conceptions of that concept. … If Hercules sits in the abortion cases, he must decide that issue and must employ his own understanding of dignity to do so. It would be silly to deny that this is a political decision, or that different judges, from different subcultures, would make it differently. Even so, it is nevertheless a very different decision from the decision whether women have, all things considered, a background right to abort their fetuses. *Hercules might think dignity an unimportant concept; if he were to attend a new constitutional convention he might vote to repeal the due process clause, or at least to amend it so as to remove any idea of dignity from its scope. He is nevertheless able to decide whether that concept, properly understood, embraces the case of abortion.*[51]

This quote suggests that even if Hercules' own convictions go one way on a topic, his job is not to use his own personal sense of right or wrong. Rather, he must attempt to discern the most sensible rendering of the concept to those for whom it has genuine meaning and importance. If this is the case, it is not clear how Hercules' judgments could differ from the judgment another judge would make on the same issue of political morality. There is a complex debate around the commensurability of fit and justification on this topic;[52] however, whether Dworkin convincingly

[50] Dworkin, 'Hard Cases' (n 37) 1095. Emphasis added.

[51] ibid 1105. Emphasis added.

[52] John Finnis, 'On Reason and Authority in "Law's Empire"' (1987) 6 Law and Philosophy 357; Ronald Dworkin, 'No Right Answer?' in Peter Hacker and Joseph Raz (eds), *Law, Morality and Society: Essays in Honour of HLA Hart* (Clarendon Press 1977); Ronald Dworkin, 'On Gaps in

responds to this issue is not relevant to this thesis. All that is important is to note that epistemically ideal judges play a vital role in both the defence and formulation of Dworkin's right answer thesis, irrespective of how defensible or consistent Dworkin's account of the epistemically ideal judge itself is.

Moore, too, has advocated a 'right answer' thesis, albeit one that is inspired by much different philosophy than Dworkin's. Indeed, in criticising Dworkin's formulation of his right answer thesis, Moore suggested not that the thesis itself was unsound, but that it was incompatible with Dworkin's conventionalism.[53] Moore's own preferred view is a 'metaphysically realist' theory of law. Moore has advanced a particular understanding of what this entails. He contends that a 'full-blooded' realist for a given class of entities (including law) will maintain: (1) that the entities in question exist, (2) that the entities are mind-independent, (3) a correspondence theory of truth whereby the predicate 'is true' derives meaning from its correspondence to some independent state of affairs, (4) a truth-conditional theory of the meaning of sentences, whereby semantics is dependent on representation of how things actually are and (5) a causal theory of the meaning of words.[54]

One of the unifying themes in both Dworkin and Moore's right answer theses is an identification of law as metaphysically *objective* in some sense. However, the extent to which Dworkin and Moore are committed to legal objectivity varies. Given that objectivity matters significantly more to Moore's theory than it does Dworkin's (per proposition (2) above from Moore's list of criteria for metaphysical realism), I return to this more abstract philosophical issue, and Moore's right answer theory, in Chapter 10.

Whatever the specifics of this objectivity, however, the general thrust of charac-

---

the Law' in Neil MacCormick and Paul Amselek (eds), *Controversies About Law's Ontology* (Edinburgh University Press 1991); Brian Bix, *Law, Language and Legal Determinacy* (Clarendon 1995) 100.

[53] Michael Moore, 'Metaphysics, Epistemology and Legal Theory' (1987) 60 S Cal L Rev 453, 483–94.

[54] Michael Moore, 'The Interpretive Turn in Modern Theory: A Turn for the Worse?' (1989) 41 Stan L Rev 871, 878–79.

terising law as objective unites the Blackstonian, Dworkinian and Moorean views. The view that law is objective is very closely linked with the doctrine of voidness *ab initio* that was identified in each of the jurisdiction studies. Indeed, the view that unconstitutionality requires voidness *ab initio* is perhaps the core common theme across the comparator jurisdictions. Chapter 9 will propose a model of legal systems that characterises such systems as incomplete (ie, they are indeterminate and have normative gaps). Chapter 10 will consider general criticisms of objectivity. If the arguments in these chapters are successful, they pose significant challenges to the void *ab initio* view.

8.3.2.3. The Role of Interpretation

The legislative nullity view will embrace something like what Solum has called the 'fixation thesis' in the context of the originalism debate in the United States.[55] This thesis is framed as follows: 'The object of constitutional interpretation is the communicative content of the constitutional text, and that content was fixed when each provision was framed and/or ratified.'[56] The term 'communicative content' here is intended to be neutral with respect to debates on what the 'meaning' of the constitutional text is meant to refer to; for example, whether the text refers to the intentions of the original authors, or the public understanding of the terms prevailing at the time.

The important part of the fixation thesis for present purposes is the idea that the meaning of the Constitution was determined at the point of framing or ratification.[57] In other words, all the legal content of the Constitution existed at the moment of framing or ratification. This is attractive for a strong believer in legislative nullity because it identifies an objective, historical base for the authority of constitutional norms. Let us call this view 'strong' legislative nullity. The proponent of this 'strong'

[55] Solum (n 32).

[56] ibid 15.

[57] The distinction between framing and ratification is a nod to another aspect of intra-originalist debate and is unimportant here.

view does not need to worry about objections that the judiciary are truly supreme because they can change the content of the constitution by interpreting it. They will say that the judiciary are not creative in this enterprise; they are more like an external observer determining the operation of a system than a theorist within that system determining how best to develop it.

Advocating legislative nullity does not compel a theorist to adopt this strong thesis. A different form of legislative nullity would embrace the 'living tree' approach to constitutional interpretation. For the sake of contrast, let us call this view 'weak' legislative nullity. This view rejects the fixation thesis, as it must hold that there is some scope for constitutional content to be added through interpretation. This sort of approach is reflected in the Canadian 'living tree' doctrine.[58] Those that subscribe to 'living tree' interpretation can still maintain that there is *some* constitutional content present at the time of enactment, but that this is not exhaustive of all the constitutional content. Thus, judges can plausibly both add to, or discover, new constitutional rules.

This thesis does not adjudicate between the strong and weak legislative nullity thesis, as these are functions of independent views on constitutional interpretation. It suffices merely to observe the relationship that obtains between these theories of constitutional interpretation and responses to the derivation question as posited by this thesis.

8.3.3. Differentiating Legislative Nullity and Judicial Intervention

One might object that the judicial intervention answer to the derivation question is not actually incompatible with the legislative nullity view, and thus they are not sufficiently analytically distinct to form separate answers to this question. This section clarifies some similarities and differences between the answers for this purpose.

[58] *Edwards v Canada (Attorney General)* [1930] AC 124 (UKPC) 136; Aileen Kavanagh, 'The Idea of a Living Constitution' (2003) 16 Can J L & Jurisprudence 55; Bradley Miller, 'Beguiled by Metaphors: The Living Tree and Originalist Constitutional Interpretation in Canada' (2009) 22 Can J L & Jurisprudence 331.

A potentially misleading commonality between the answers is that both of them alter the legal community's understanding of what the true state of the past law was.[59] The means by which this is achieved in each answer, however, is different, and is the nub of why the answers are incompatible. Under the legislative nullity view, the alteration of past law is the result of a non-judicial power not being exercised; it is a failure of legislation. This is not something that is capable of revision or redemption by the judiciary alone. It is thus incompatible with a view of the judge as having any law-making power.

With regard to the judicial intervention answer, the justification for the decision is some pre-existing legal standard that is relevant to deciding the case and articulating the unconstitutionality. This true state of the law could reflect that a statute previously thought to be valid is, in fact, invalid. This would be similar to what the legislative nullity answer achieves. However, precisely because the discovery of these standards is something that is accomplished by the *judiciary*, the alteration of the past law produced by the judicial intervention answer is more qualified. Later judges can always argue that they have better insight into what the law 'truly requires' than their forebears. Conversely, when a judge observes that something was not properly legislated, the act of so observing takes the matter out of the judge's hands; if the legislator failed to legislate because of constitutional norms that are, *ex hypothesi* on the legislative nullity view, not themselves manipulated by the judiciary, then there cannot truly be scope for revision of unconstitutionality by the judiciary once it is discovered.

This difference reflects a broader point that the legislative nullity view is bound up with Blackstonian ideas about fixed interpretation and objective meaning in law. On the other hand, the judicial intervention view is a natural companion to a realist model of interpretation where the law has a more fluid and dynamic character over

[59] I am deliberately avoiding terming this 'retrospectivity', as neither answer entails that a particular law be retrospectively effective. The legislative nullity and judicial intervention views 'reveal' the true juridical status of laws in the past. They are not about making entirely new law and having it apply to facts pre-dating its enactment, which is what I understand by legal retrospectivity.

time. These points are often not borne out clearly in judicial statement however. In the context of delivering judgment in a case, and thus responding to the exigencies of that case, judges will understandably react more to moral impulse than technical and theoretical clarity. This means that these models are often deployed to justify a result rather than as part of a clear and overarching understanding of the judicial role. This, combined with the point observed above regarding the simulation of retrospectivity, can make the models difficult to disentangle from one another in case law.

This distinction is more clearly observable in some doctrinal disputes that have arisen in some jurisdictions. The most obvious example is the 'revival' of dead statutes.[60] That is, the restoration of an unconstitutional statute pursuant to the reason for the original unconstitutionality being displaced. This practice is only compatible with the judicial intervention view. On the legislative nullity view, the judiciary may only find a statute unconstitutional. It cannot revisit that consequence later even if it revises its opinion, or even if it revises the substantive constitutional norm that made the legislation unconstitutional in the first place.[61] Once a failure of legislation is identified, the onus falls on the *legislature* to take further action.[62] By contrast, since the judiciary partly establish the unconstitutionality themselves on the judicial interview, they can disestablish it in later cases and immediately reinstate the previously stricken law.

---

[60] Recall that, as described in chapter 3, this possibility was observed through some older US authorities: *Legal Tender Cases* 79 US 457 (1870); *Leisy v Hardin* 135 US 100 (1890); *In Re Rahrer* 140 US 545 (1891); *Adkins v Children's Hospital* 561 US 525 (1923); *West Coast Hotel Co v Parrish* 300 US 379 (1937).

[61] Recall the difficulties in parsing how anything could amend a nullity: *Saghir Ahmad v State of UP* [1954] INSC 89, [1955] SCR 707; *An Blascaod Mór Teoranta v Commissioners of Public Works (No 3)* [1998] IEHC 38.

[62] It should be noted that a practice that superficially looks like revival may still occur under the legislative nullity view. This would arise where a statute is 'mistakenly' stricken according to a flawed constitutional interpretation. This is how the legislative nullity view will parse constitutional revisions: by presenting them as mistakes. However, the difference is that under the legislative nullity view the statute does not 'revive' because it was never dead. It was, on that view, a mistake to treat the statute as ever having been dead in the first place. The legislative nullity view does not allow for norms in abeyance.

Another area of constitutional doctrine that is impacted by, or reflects, the difference in answers to the derivation question is the idea of 'creeping unconstitutionality'. This issue has been given at least some consideration in both Ireland and India. Creeping unconstitutionality refers to the problem of whether legislation can 'become' unconstitutional over the passage of time, either because of factual changes[63] or evolution in moral values.[64]

Most relevantly for the argument in this thesis, there is some authority to suggest that what matters is whether the legislation was constitutional at the time of its enactment, not whether it remains compliant with changing constitutional trends. In *Browne v Attorney General*, the Irish High Court held that:

> If the legislation was constitutional at the time of its enactment it would seem impossible to contend that it was ever *ultra vires* the Oireachtas … to enact the legislation. In these circumstances it seems to me that the challenge would have to be made on the basis that the Oireachtas by failing to repeal or amend the [impugned act] … failed to vindicate the constitutional rights of the plaintiffs.[65]

By contrast, however, it seems to be the case that unconstitutionality can 'creep' in the Indian context. As Basu remarks:

> The law … may be constitutional when enacted but with the passage of time, the same may be held to be unconstitutional in view of the changed situation. … If a provision was not unconstitutional on the day on which it was enacted or the Constitution came into force, by reason of the facts emerging out thereafter, the same may be rendered unconstitutional such as on the ground of gender equality.[66]

The point of relevance for present purposes is this: it is difficult to see how unconstitutionality can 'creep' consistently with the legislative nullity view. On the stipulations of that theory, only a formal amendment process could further add or repeal constitutional norms. Assuming the truth of that theory, a piece of legislation should indeed, as the Irish High Court had it, be constitutional or unconstitutional by reference to the constitutional norms that existed at the time it was enacted. Of

---

[63] *Blake v Attorney General* [1982] IR 117 (SC).

[64] *McGee v Attorney General* [1974] IR 284 (SC).

[65] *Browne v Attorney General* [1991] 2 IR 58 (HC) 69–70.

[66] Durga Das Basu, *Commentary on the Constitution of India* (8th edn, LexisNexis 2011) 61.

course, the judiciary may play a role in identifying which constitutional norms prevailed at the time of enactment, but they do not create or constitute these norms. By contrast, unconstitutionality could creep on the judicial intervention view. This analysis would be consistent with the analysis in chapters 5 and 6, which suggest that the Irish courts tend towards a legislative nullity view, whereas the Indian courts tend towards judicial intervention.

8.4. THE EFFECT OF UNCONSTITUTIONALITY

> *What does unconstitutionality do? In what way does unconstitutionality modify legal norms?*

The preceding section addressed the question of which power is constitutive of unconstitutionality. This section seeks to answer a separate question: what does unconstitutionality, itself, achieve? Unconstitutionality is an operation that modifies a legal norm, but what does it modify? I canvass two major approaches to this question: unconstitutionality either *invalidates* a legal norm, or it renders it *inapplicable*.

These properties cannot be taken in total isolation. This is due to a significant degree of conceptual confusion around the formal properties of legal norms, with the same or similar terms being used by different theorists to refer to different properties. I will provide an account of the following properties of legal norms: validity, existence, and applicability. Two of these concepts—existence and validity—are already quite familiar in legal philosophy due primarily to the influence of Kelsen's work. Indeed, they are often treated as coextensive. For this reason, these two more 'classic' properties are analysed together.[67] The third property, applicability, is less

[67] These are often analysed with a third property: the *efficacy* of legal norms. I do not consider efficacy here because the usual understanding is that efficacy is a factual measure of the degree to which legal norms actually prompt adjustments in the behaviour of the norm population: Hans Kelsen, *General Theory of Law and State* (Anders Wedberg tr, Harvard University Press 1945) 39–40. This is not relevant for the effects question. Unconstitutionality does not, and cannot, modify the efficacy of a legal norm. Munzer has illustrated why efficacy should not be equated with existence:

> A [duty imposing] rule may not succeed in bringing about the appropriate or intended state of affairs; that is, action-conformity … may nearly always be lacking. The rule would then be inefficacious. But

well attended to and so will be analysed separately to see how it contrasts with the other two and forms an analytically distinct property of legal norms, and one that is crucial to understanding the consequences of unconstitutionality.

8.4.1. Validity

We must begin first by defining what legal validity and applicability mean. Let us start with validity. The validity of a norm is a function of its system membership. Valid legal norms are those norms that are properly promulgated within a system. Consequently, invalid norms are those that are rejected as non-members of the system because of contradictions with higher norms within that system. Occasionally, however, validity is taken to establish something more than just this membership idea. To understand how, let us begin with Kelsen's idea that validity is the 'specific existence' of legal norms.[68] This immediately sets up a strong relationship between the validity of a norm and its existence to the point where, for Kelsen, the invalidation and repeal of a legal norm become indistinguishable.[69] Unfortunately, Kelsen's use of the terms validity and existence is somewhat ambiguous. In particular, he fails to sufficiently distinguish between a normative concept of validity—validity as binding force—and a descriptive concept of validity as membership of a system.[70] This strong link between validity and existence, and the equivocation of the term 'validity', are the theoretical data that provide for the orthodox 'void *ab initio*' theory of unconstitutionality that has been taken up in many of the courts studied in this thesis.

> *if its violators were invariably or almost invariably punished, there would be no ground for denying the existence of that rule.*

Stephen Munzer, *Legal Validity* (Martinus Njihoff 1972) 35. Emphasis added.

[68] Kelsen, *PTL* (n 9) 213; Dick Ruiter, 'Legal Validity Qua Specific Mode of Existence' (1997) 16 Law & Philosophy 579.

[69] '[A] conception of legal validity as a mode of existence brings with it the idea that there is no fundamental difference between invalidating and repealing a legal norm.' Dick Ruiter, 'Legal Validity Qua Specific Mode of Existence' (1997) 16 Law & Philosophy 579, 480.

[70] Eugenio Bulygin, 'An Antinomy in Kelsen's Pure Theory of Law' (1990) 3 Ratio Juris 29, 36–41.

As mentioned above, for Kelsen the validity of a legal norm is its 'specific existence'. Thus, a legal norm 'exists' to the extent that it is valid. On Kelsen's theory, legal validity is a relative notion insofar as a lower norm is valid only relative to a higher one.[71] Because law obtains its validity in this way, it must always be a member of an ordered set of legal norms, else each legal norm would have no hierarchical source from which to derive validity. That the set of legal norms must be *ordered* determines that it is a legal *system* and not merely a smattering of legal norms or normative statements.[72] This fits well with the observations on legal systems that will be made in chapter 9.

Kelsen is not alone in maintaining a strong link between validity and existence, though it has its origins in his work. Other modern legal positivists expressly endorse the claim that validity and existence are co-extensive.[73] The claims of modern legal positivism might be summarised by the following two theses: *(i)* the matter of the existence of a legal system turns on its effectiveness; and *(ii)* the matter of existence of any law depends on the effectiveness of its parent system and its source within that system.[74]

For present purposes in considering unconstitutionality, thesis *(ii)* is of the most significance. The word 'source' in that claim restates the Kelsenian point that for legal positivists the validity of a law turns on a particular relative relationship that obtains between one legal norm and another hierarchically ordered legal norm within the same legal system. This is particularly problematic in the case of laws that are invalid. On the one hand, these norms have no source in the legal system and so do not 'exist' within that system; but, on the other, as Waluchow has observed, a law

---

[71] Natural law theories, by contrast, might posit a more absolute criterion of legal validity albeit one that is probably still relative to some moral norm, rather than a higher-order legal one.

[72] Navarro and Rodríguez (n 8) 119.

[73] Leslie Green 'Legal Positivism', *The Stanford Encyclopedia of Philosophy* (Fall edn, 2009) ⟨http://plato.stanford.edu/archives/fall2009/entries/legal-positivism/⟩ accessed 5th December 2014; Joseph Raz, 'Legal Validity' in Joseph Raz, *The Authority of Law: Essays on Law and Morality* (2nd edn, OUP 2009) 146.

[74] Both these theses are taken from Joseph Raz, 'Legal Validity' in Joseph Raz, *The Authority of Law: Essays on Law and Morality* (2nd edn, OUP 2009) 152.

that is 'accepted and practiced as valid law does seem to exist even if I ought not to obey it'.[75] This is effectively the issue of unconstitutional laws that are deemed invalid.

The first step to disentangling this knot rests in the distinction between normative and descriptive validity briefly mentioned above. This distinction is derived from von Wright's highly influential work on the formal properties of norms.[76] Von Wright identifies two meanings of validity: the binding force of a norm, which is a normative concept,[77] and the promulgation of a norm in an act that was, itself, legal. This latter concept is a descriptive concept. Let us call these N-validity and D-validity here for simplicity.

Von Wright suggests that to identify the existence (itself a descriptive concept, and one that is not necessarily the same as D-validity) of a norm with the N-validity of another norm results in an infinite regress.[78] This is because the N-validity of a child norm is properly derived from the D-validity of its parent norm. This gives priority to the descriptive sense of validity over the normative sense. If the distinction between these senses of validity is collapsed, as it would be in the case of co-extension between validity and existence, a significant difference between normative and descriptive derivation becomes obscured. This can be seen by hypothesising a purely normative scheme of legal validity. If the N-validity of every norm was derived from the N-validity of a higher norm, then it becomes impossible to determine the root of validity of the legal system because one can never stop asking: 'from which valid norm does this norm derive validity?'[79] In this hypothetical, because the bindingness of a norm is derived from the bindingness of a parent norm, the

---

[75] Wil Waluchow, 'Four Concepts of Legal Validity' in Matthew Adler and Kenneth Himma (eds), *The Rule of Recognition and the US Constitution* (OUP 2009) 138.

[76] Georg Henrik von Wright, *Norm and Action: A Logical Enquiry* (Routledge & Kegan Paul 1963).

[77] ibid 195–96.

[78] von Wright (n 76) 196–97; A separate regress argument for legal validity can be found in: Oliver Black, 'Legal Validity and the Infinite Regress' (1996) 15 Law and Philosophy 339.

[79] An instantiation of this problem is Kelsen's theoretical postulate—the *Grundnorm*—which is designed to avoid this very circularity.

derivation cannot end at a descriptive proposition; it must continuously seek out higher-order binding norms.

There is a way around this difficulty. Validity is a relational property, but it must relate the *(N-)validity* of one norm (the subordinate norm) to the *existence* of a superordinate norm. This avoids the potential circularity problem by ending it with a concrete descriptive proposition (the existence of the apex norm of the legal system) but this resolution requires that the existence of a legal norm is ontologically prior to validity, not co-extensive with it. This observation aside, it is true to say that in much legal thought Kelsen's view has become at least tacitly accepted, and validity and existence are often treated as though they are co-extensive. Notwithstanding its vaunted status as a theoretical postulate, this supposition leads to difficulty in practice when dealing with unconstitutional norms.

In a sophisticated attempt to get to the root of the problem of the equivalence between legal validity and existence, Waluchow suggests that there are four different concepts of legal validity:[80]

(1) *Legal validity as existence*: R is officially accepted and practised in legal system L as a norm that fully satisfies all systemic criteria of legal validity (both pedigree- and merit-based) included within L's rule(s) of recognition.

(2) *Systemic Validity*: R is officially accepted and practised in legal system L as a norm that fully satisfies all systemic criteria of legal validity (both pedigree- and merit-based) included within L's rule(s) of recognition, and does, as a matter of objective fact, satisfy all such systemic criteria of validity.

(3) *Systemic Moral Validity*: R is officially accepted and practised in legal system L as a norm that fully satisfies all systemic criteria of legal validity (both pedigree- and merit-based) included within L's rule(s) of recognition; does, as a matter of objective fact, satisfy all such systemic criteria of validity; and 'has the normative consequences [it] purport[s] to have' because it is the product

[80] Waluchow, 'Four Concepts of Legal Validity' (n 75) 140.

of a legal system that (a) fulfils 'the need to have effective law' and (b) issues from 'a justified authority.'

(4) *Moral Validity*: R is morally justified on its own terms, that is, independently of its membership in L.

Of particular interest for present purposes is the distinction between senses (1) and (2), which it seems is meant to eliminate some subjectivity in legal validity. The distinction between these senses is meant to capture a certain intuition with respect to unconstitutional norms that have, in fact, been relied upon. All that matters for sense (1) is that the relevant officials *treat* the norm as though it were, say, passed by the legislature. However, this seems overbroad *qua* standard of validity. As there is no requirement that R in sense (1) need have anything to do with the systemic criteria of legal validity, it does not seem to stop an example of a legal official (say, a judge) taking some extra-legal standard and using it as a justification for legal decisions.[81] Certain extra-legal norms can be accepted in legal systems, and might even regularly feature as parts of legal judgments, but this does not make them legally valid. Of course, these would fail to be systemically valid, but this does not seem to strictly matter under the terms of sense (1).

Sense (1) is designed to allow for unconstitutional legal norms (those that have no pedigree-based criterion of validity in a system) to be 'valid' in some sense. It is questionable, however, whether Waluchow's sense (1) of validity goes further than this: for example, would it allow for wilful self-deception or counter-factual insistence by legal officials? It seems odd to predicate a sense of legal validity entirely on something as subjective as a collective consciousness from such an ambiguous group as 'legal officials'. As a sense of legal validity, therefore, (1) seems to suffer from some defects. Nevertheless, Waluchow is still capturing an important property

[81] I am not claiming that such norms cannot be legally relevant, however. Non-legal norms play important roles in legal decision-making. Take norms of logic, for example. For some statements as to the importance of norms of logic in law, see: John Gardner, 'Legal Positivism: $5\frac{1}{2}$ Myths' in John Gardner, *Law as a Leap of Faith* (OUP 2012) 212, fn 29; Desmond Clarke, 'Judicial Reasoning: Logic, Authority, and the Rule of Law in Irish Courts' (2011) 46 Ir Jur 152.

of legal norms here; I suggest that sense (1) of validity captures what it means for a norm to be legally *applicable*, but not legally *valid*. It is thus better to view sense (1) of validity above as capturing an important property of legal norms but not a property that has anything to do with validity.

The net point of the discussion of validity is that cashing out a distinction that can preserve the justificatory force of a law—in the sense that it can still act as a good ground for institutional action—while at the same time condemning it for being a non-member of the legal system is difficult to do in pure validity terms. For this reason, I now turn to consider the separate property of applicability.

8.4.2. Applicability

Some conceptual knots in which legal theory becomes tied when analysing validity and existence can be unwound by analysing the related, but distinct, concept of applicability. Although there are traces of this concept in the writings of Raz,[82] Munzer,[83] and Hart,[84] the most developed account of the applicability of legal norms—particularly how applicability may be a property of legal norms that is independent of their validity—is to be found in Bulygin's work, as developed by others.[85]

Validity, as we have seen, is a function of the *origin* and the *content* of legal norms. It is a question of genealogy (who promulgated the norm) and coherence (does the norm cohere with higher-order norms of the same system). Applicability, by contrast, is primarily a function of the *scope* and *justificatory power* of legal

[82] Joseph Raz, 'The Institutional Nature of Law' in Joseph Raz, *The Authority of Law: Essays on Law and Morality* (2nd edn, OUP 2009) 119–20.

[83] Munzer, *Legal Validity* (n 67); Munzer, 'Validity and Legal Conflicts' (n 20) 1156–58-.

[84] Hart, *CL* (n 19) 261–62.

[85] Eugenio Bulygin, 'The Problem of Legal Validity in Kelsen's Pure Theory of Law' in Carlos Bernal and others (eds), *Essays in Legal Philosophy* (OUP 2015) 132–36; Eugenio Bulygin, 'The Problem of Legal Validity in Kelsen's Pure Theory of Law' in Carlos Bernal and others (eds), *Essays in Legal Philosophy* (OUP 2015); Navarro and Rodríguez (n 8); Pablo Navarro and others, 'Applicability of Legal Norms' (2004) 17 Can J L & Jurisprudence 337; Pablo Navarro and José Juan Moreso, 'Applicability and Effectiveness of Legal Norms' (1997) 16 Law & Philosophy 201.

norms.[86] It relates to the types of case covered by the norm, and those in which the norm can justify legal results; it describes a norm's ability to regulate a case and uphold a (legal) conclusion in that case. This ability to justify legal conclusions must be granted by some other application norm of the system. The difference between the scope and justificatory power functions of applicability is reflected further in the distinction between 'internal' and 'external' applicability, which I describe below.

For now, recall that part of the problem of invalid norms is that they seem to founder on a peculiar paradox. Invalid norms are norms that are not members of the legal system, and yet they still claim normative validity, either presently or for some duration of time that is now past. Through the lens of legal validity, as was argued above, this looks like a non-existent norm (a norm that is not a member of the legal system) pretending to a validity (bindingness) claim. But how can something that does not exist be binding? Bulygin suggests that one way to resolve this problem would be to make the existence of norms a function of applicability rather than validity:

> [A] derogated norm—provided it is still applicable to certain cases—has not been eliminated from the system. It continues to be a member of the system although its applicability has been restricted by derogation to a more limited range of cases; for example, it might no longer be applicable to future cases. Thus, derogation only limits the applicability of a norm but does not deprive it of its existence in a system because it does not remove the norm altogether from the system.
>
> [...]
>
> This suggestion amounts to shaping the concept of the existence of a norm … in such a way that existence become a function of applicability: in order to know what norms are members of a given system of law, one must know what norms are applicable. Then, however, we must be prepared to say that the existence of a norm does not begin from the moment of its promulgation but instead from the moment of its applicability.[87]

This may seem attractive at first blush, but Bulygin himself rejects this view.[88] To understand why, consider how it would apply to establishing the existence of

[86] Pablo Navarro and others, 'Applicability of Legal Norms' (2004) 17 Can J L & Jurisprudence 337, 337.

[87] Eugenio Bulygin, 'Time and Validity' in Carlos Bernal and others (eds), *Essays in Legal Philosophy* (OUP 2015) 184.

[88] ibid 185.

a contract. If A and B contract that A will ϕ tomorrow, no obligation arises until tomorrow. It would be strange to suggest that the contract does not exist today, but this would be implied by an applicability-as-existence view. This view would also bear the strange implication that wills do not 'exist' until after the testator's death.

In further support of this division between applicability and existence, Bulygin also notes that the criteria for the applicability of legal norms are to be found in the law. Specifically, they are part of the legal system that corresponds to the present time.[89] This means that one must identify this legal system without reference to the applicability criteria, else one would be supplying a circular definition for these criteria. Thus, just as validity is not coextensive with existence, neither is applicability. This establishes, at least, that existence is a standalone property of legal norms that is not fully entailed by either of the other properties under discussion.

The analysis above suggests that as long as a legal norm is a member of *some* legal system, then it can be applicable even outside of that system.[90] Munzer also makes a similar concession.[91] How, then, can we parse unconstitutional norms in the language of inapplicability? Navarro and Rodríguez offer the following reconstruction:

> An unconstitutional norm is an invalid norm, because it does not meet the appropriate systematic relations with other norms of a legal system. Nevertheless, it may happen that the competent authorities empowered to control their constitutional validity, such as a constitutional court, wrongly declare them valid. To deal with these invalid norms, we must bear in mind the distinction between final and infallible judicial decisions, because *the constitutionality of a legal norm does not depend on what the constitutional court decides*. […] However, *the final decision of the constitutional court determines the applicability of those norms*. If the court wrongly says that a certain norm is constitutional, that norm will be applicable, although invalid in the system.[92]

---

[89] ibid 177.

[90] It is important here to note that Bulygin's model of legal systems is built of several 'legal systems' (which are fixed to individual points in time) being linked together in a 'legal order'. Thus, for Bulygin, unconstitutional norms can exist in a legal system at some point in the past and still, therefore, be a part of the legal order. This idea of legal systems being fixed to points in time (rather than persisting over durations of time) is also shared by Raz: Joseph Raz, *The Concept of a Legal System: An Introduction to the Theory of Legal System* (Clarendon Press 1980).
The legal theory around legal systems is considered in greater detail in chapter 9.

[91] Stephen Munzer, 'Retroactive Law' (1977) 6 J Legal Stud 373, 386–87.

[92] Navarro and Rodríguez (n 8) 134. Emphasis added.

This can occur where a sufficiently high-ranking court positively affirms a mistake in law.[93] Doctrines such as *res judicata* can operate to make such a mistaken view applicable. Similarly, in cases where a norm's validity is positively rejected by the court, *res judicata* can confer applicability on the invalid norm. In other words: both positive mistakes of law and rejections of legal views as mistaken can be retrospectively and prospectively applicable. It is for this reason that, for example, prisoners convicted on foot of an unconstitutional offence are not all summarily released from their custodial sentences once the unconstitutionality is declared. At best, some may be able to initiate *habeas corpus* proceedings; however, by default the unconstitutional offence is still applicable to their cases because those cases are considered settled and *res judicata*.

One might think, on the basis of the foregoing, that applicability does not bear meaningfully on the voidness *ab initio* problem, which expressly contends that the impugned norm was never a member of *any* legal system within the legal order.[94] This is still not a difficulty, however. Applicability applies to a wider class of norms than just *legal* norms. All the void *ab initio* theory can establish (at best) is that the norm is not legally valid in any legal system that is part of a jurisdiction's legal order. Voidness *ab initio* would thus only control applicability questions if legal validity were necessary for legal applicability. I argue that this suggestion does not hold; legal validity is *not* a precondition of applicability. To cash this claim out, I turn now to the analysis of applicability undertaken by Navarro and Moreso.

---

[93] See also: Kenneth Himma, 'Final Authority to Bind With Moral Mistakes: On the Explanatory Potential of Inclusive Legal Positivism' (2005) 24 Law & Philosophy 1.

[94] I describe a theory of legal systems that further explains these terms in chapter 9. For now it suffices to note that 'legal system' here describes a collection of legal norms *relative to a particular point in time*, and 'legal order' refers to a sequence of these legal systems that relates them to one another and provides for continuity between them. The 'legal order' is thus what we are really referring to when we discuss entities such as the 'Irish legal system'.

8.4.3. Invalidity does not Imply Inapplicability

It was argued above that neither validity nor applicability are coextensive with a legal norm's existence. This establishes that existence is some separate property, but still leaves open the possibility that validity and applicability are coextensive with each other. In support of separating out membership (validity) and applicability, one might observe that not all items that are recognised as applicable within a system are themselves members of that same system, as evidenced by the rules of private international law, *inter alia*.[95] All legal norms must belong to *a* legal system; there are no free-floating legal norms. However, merely establishing that all legally valid norms must be a member of some legal system does not establish that all the norms applicable within a given legal system are member-norms of that particular legal system. The set of norms *applicable* in a legal system may not necessarily be identical to the set of norms that are *valid* in that system.

To see how this can be so, it is useful to consider examples of norms that might be (1) invalid but applicable, and (2) valid but inapplicable. The most obvious examples of norms that are invalid but applicable are the norms of foreign legal systems, through application of the rules of conflict of laws. From the perspective of the domestic legal system, foreign legal norms are non-members of the system. That is, they are invalid, from the perspective of that system. Notwithstanding this invalidity, they may still be deemed appropriate to apply to certain cases. Additionally, in some cases, law that has been repealed may continue to be applicable to sets of facts that arose before that law was repealed. This will often be provided for expressly by repealing legislation, which will apply a proviso for the limited continued application of the old law where relevant. The key point to note is that neither foreign law nor law that has been repealed are valid members of the domestic legal system. Again, as validity is determined by membership, this effectively means that these laws are invalid in that system. Nevertheless, these laws may still be applicable in appropriate circumstances.

[95] Raz, 'Legal Validity' (n 74) 148–49.

To see how norms can be valid but inapplicable, consider legislative norms that are subject to delayed commencement, as briefly discussed in chapter 2.[96] That is, those legislative norms that have been assented to, but have provided for some other date for the law to enter into force (become applicable). Consider also norms that have fallen into desuetude.[97] There is increasing acknowledgement that complex social processes can form conventions that count as law.[98] My concern here is not the complex process by which a norm might fall into desuetude; there is a difficult symbiosis here whereby a norm can become inapplicable because it is unapplied. I do not intend to provide an account of this process. The point for present purposes is that once a norm is recognised, or tacitly understood, as having lapsed into desuetude, it may become inapplicable. That is, judges may view it as incapable of justifying legal results in cases.

A further example might be contractual norms. If a person A contracts to buy 30 widgets from B, he is under a legal obligation to reimburse B. The norms governing

---

[96] In civil law this is analogous to *vacatio legis*, the period between when a law is enacted and when it becomes applicable. Provisions of *vacatio legis* are quite common in many instances. Bulygin gives the example of the Argentine Criminal Code. This was enacted on the 29 October 1921. However, article 303 of the Code provided that it would only be applicable six months later starting from the 29 April 1922. See: Bulygin, 'Time and Validity' (n 87) 174–75.

[97] The extent to which desuetude applies varies between jurisdictions and even between types of law. It has been analysed in Ireland (Gerard Hogan, 'Statutory interpretation – the doctrine of desuetude' (1987) 9 DULJ 136; Maebh Harding, 'The Curious Incident of the Marriage Act (No 2) 1537 and the Irish Statute Book' (2012) 32 Legal Stud 78) and the United States (Linda Rodgers and William Rodgers, 'Desuetude as a Defence' (1966) 52 Iowa L Rev 1; Mark Henriques, 'Desuetude and Declaratory Judgment: A New Challenge to Obsolete Laws' (1990) 76 Va L Rev 1057). Its application to achieve constitutional amendment has also been considered extensively by Albert: Richard Albert, 'Constitutional Amendment by Constitutional Desuetude' (2014) 62 Am J Comp Law 641; Richard Albert, 'Constitutional Disuse or Desuetude: The Case of Article V' (2014) 94 Boston Uni L Rev 1029; Richard Albert, 'The Desuetude of the Notwithstanding Clause – And How to Revive It' in Emmett Macfarlane (ed), *Policy Change, Courts, and the Canadian Constitution* (University of Toronto Press 2018).
Interestingly, the Supreme Court of India has positively accepted the doctrine: *State of Kerala v NM Thomas* [1995] INSC 178, [1995] SCC 3 434. The anti-desuetude trend in the common law tradition is rooted in writers considering that desuetude must involve the *repeal* of unused statutes. This is the only way those writers can rationalise a statute becoming ineffective. Once the distinction between validity and applicability is observed—as through the example of delayed commencement given in the text—it is much more plausible to see desuetude as rendering a statute inapplicable rather than asserting that it has become invalid.

[98] See, for example: Oran Doyle, 'Conventional Constitutional Law' (2015) 38 DULJ 311.

contract formation generally are valid norms of the system, but the specific obligation of A to pay B is not a valid norm of the system. Recall that validity requires a norm to be a *member* of the legal system. It would be odd to say that A's obligation to pay B is a member of the system (that is, recognised by the rule of recognition). But we still correctly think that A must pay B, or a court will be justified in finding against him and imposing a remedy. So, the norm for A to pay B is applicable to a case concerning A and B. A further interesting feature here is that the reason a third party, C, is not bound is by operation of the doctrine of privity of contract. Privity is, on this analysis, a rule governing the applicability criteria of contractual relations. Specifically, it maintains that (subject to certain narrow exceptions) contractual norms are only applicable to the original contracting parties that agreed them.

Now that it has been observed that there are at least some cases in which validity and applicability *can* come apart, the following sections will present a more conceptual analysis of the separation between validity and applicability. The key idea is that there is *no connection of necessity or equivalence* between validity and applicability. It is very important to stress that this is different from saying that there is no connection between applicability and validity at all. There are lots of contingent, that is, unnecessary but actual, connections that exist between validity and applicability. Thus, in most cases, there will be no need to analyse them separately. However, unconstitutionality is by its nature an aberration in the law. Something has gone wrong when a law is declared unconstitutional. It is in cases such as this, as well as the conflicts, desuetude, etc cases mentioned above that distinguishing validity and applicability may assist in achieving conceptual clarity.

8.4.3.1. D-Validity ≠ Applicability

Navarro and Moreso observe that membership of a norm in a legal system is a sufficient (but unnecessary) condition for what they term internal applicability and membership of a norm in a legal system is both unnecessary and insufficient for what they term external applicability. It is worth reproducing the definitions of each

in full:

> *External Applicability:* A norm N_i is externally applicable at the moment t to a certain individual case c, which is an instance of the generic case C, if and only if a norm N_j, belonging to a legal system LS at t, prescribes (obliges or empowers) the application of N_i to the individual cases that are instances of C.[99]
>
> *Internal Applicability:* A norm N_i is internally applicable at moment t to an individual case c if and only if c is an instance of a generic case C, and C is defined by the spatial, material, personal and temporal spheres of validity of N_i.[100]

Internal applicability is a relatively simple idea. Norms cover generic cases ('C'). Individual cases ('c') may fall into these cases, and so are governed by the norm. Consider a hypothetical mandatory norm of etiquette such as 'upon meeting someone new for the first time, you must shake their hand'. For an individual case to be governed by this norm, two individuals must be meeting *for the first time*. It is therefore internally inapplicable to any meetings between two individuals for the second, third, etc time. This type of applicability is termed 'internal' because it is a relationship 'within' the norm, so to speak. It is a relationship between a norm and its scope.

External applicability builds somewhat on this idea. A norm cannot be externally applicable without being internally applicable. However, external applicability also requires that there is a norm of the system that makes it the case that a judge is obliged or empowered to use the internally applicable norm in a case to reach a result or justify a consequence. To build on the example used when explaining internal applicability, suppose that the handshake norm came from the 'Etiquette Act', as enacted by the legislative power governing community X. Courts in community X might use an applicability norm like 'all valid, non-derogated statutory norms that are internally applicable to cases are apt to justify legal results and consequences in those cases'. In this case, the handshake norm would be both internally applicable and externally applicable. This type of applicability is termed 'external' because it will depend on factors that exist outside of the norm itself. The external applicab-

[99] Pablo Navarro and José Juan Moreso, 'Applicability and Effectiveness of Legal Norms' (1997) 16 Law & Philosophy 201, 203.

[100] ibid 206.

ility of a norm will depend on another legal norm setting down some criterion or standard for applicability.[101]

The etiquette example can also be used to show how internal and external applicability can come apart. Imagine community Y, a neighbouring community with a separate legislature and statute book. The handshake norm is equally internally applicable to first meetings between individuals within that community. As long as they are meeting for the first time, their conduct fits within the generic category of case covered by the handshake norm. However, if the courts of community Y do not use this norm as a basis for the sanction or reprimand of individuals who do not comply with it, then it is externally inapplicable.

Navarro and Moreso observe that the type of applicability that an unconstitutional norm may lack is external applicability, but it will retain internal applicability:

> [S]uppose that a norm N does not belong to a legal system LS because it fails to meet a constitutional test of validity. N is invalid, but it is internally applicable, i.e. it regulates some actions within its spheres of validity. A quite different problem is whether we must obey an invalid norm, e.g. an unconstitutional norm, although it must be realized that a discussion on the obedience of unconstitutional norms makes sense precisely because these norms regulate behaviour.[102]

There is an important insight in this quote; unconstitutional norms are still norms as they retain internal applicability. We must therefore qualify any statement to the effect that 'unconstitutional norms do not exist (are void *ab initio*)' with the proviso that they do not exist in a certain *legal* way. Moreover, the justificatory norm referred to above (N_j) does not need to be a validity-regulating rule of the system. As Pino has observed, 'sometimes judges have a (legal) duty to apply non-valid law, as well as they can have a (legal) duty not to apply valid law'.[103] There could be many applicability norms in a legal system, and these will not always comprise the same

[101] Navarro and Rodríguez (n 8) 135.

[102] Navarro and Moreso (n 99) 210.

[103] Pino uses the phrase 'non-valid law' here in a way that encompasses my usage of both that term and 'invalid law'. Giorgio Pino, 'Farewell to the Rule of Recognition?' (2011) 5 Problema 265, 278.

set of rules as the validity-governing norms. A constitutional test of *validity* may not be the same as a constitutional test of *applicability*. To give an example of this, Navarro and Rodríguez posit the following norm:

> If the law in force at the time [a] crime was committed is different from the law in force at the time of sentence or service thereof, the law more favorable to the accused shall be applied.[104]

It must be assumed that this norm is directed to laws declaring sanctions for criminal conduct. If such a norm were applied, it would be easy to see how validity and applicability may come apart. Say N_1 is the law in force at the time an individual commits the relevant offence, and N_1 provides for a mandatory custodial sentence of 10 years. Subsequently, before final hearing of the trial, N_2 repeals N_1 and provides for a mandatory custodial sentence of 15 years. According to the norm given above, N_1 would still be applicable to this individual's case, even though it has been repealed and thus is no longer a member of the legal system (ie, it is invalid).[105]

However, it must be acknowledged that there is often a connection between validity and applicability. It is important not to mischaracterise this connection. Navarro and Rodríguez outline a view that is quite easy to slide into, which they attribute to Munzer:

> According to Munzer, legal norms must be applied (are externally applicable) to all the cases and only the cases they regulate (those in which they are internally applicable), and only valid norms that regulate a case (those that are internally applicable) should be applied to it (are externally applicable to it). This is tantamount to identifying external applicability with internal applicability plus validity, an idea that seems to capture an important number of ordinary intuitions related to the task of applying the law.[106]

---

[104] Navarro and Rodríguez (n 8) 134.

[105] In a manner not dissimilar to the analysis in the text here, the Supreme Court of Canada held in *R v Wigman* [1987] 1 SCR 246 (SCC) 257 that cases could not be reopened even where they were decided on the basis of invalidated statutes. This seems to require an acknowledgement that invalid statutes can still be applicable to certain cases (they justified, and continue to justify, the legal result in those cases).

[106] Navarro and Rodríguez (n 8) 129–30. Citing: Munzer, 'Validity and Legal Conflicts' (n 20) 1149–50.

The fact that this view captures important ordinary intuitions, and that it produces the right result in most cases, does not mean that it is actually correct in describing the relationship between applicability and validity. As mentioned above, in more exceptional or unusual cases this close alignment begins to break down. Consider cases in which there is a conflict between two norms, both of which are valid and both of which are internally applicable to a case: two statutory provisions, for example. In such cases, the legal system will apply rules to resolve the conflict; in the given example, the doctrine of implied repeal (*leges posteriores priores contrarias abrogant*) will require a court to apply the most recently enacted provision.[107] Examples such as these show that 'the intuitive idea of a legal system as a set of valid norms must be distinguished from *systems of applicable norms*, or *applicable systems*'.[108]

8.4.3.2. N-Validity ≠ Applicability

The relationship between bindingness and applicability iss one of mutual exclusivity, at least on the understanding of bindingness advocated by Kelsen.[109] For Kelsen, the bindingness of norms seems to be dependent on their being efficacious. Efficacy here refers to the idea that the norm's target population actually uses the norm as a guide to conduct. Consider how Kelsen parses desuetude:

> [A] norm may lose its validity by never being applied or obeyed—by so-called *desuetude*. *Desuetudo* may be described as negative custom, and its essential function is to abolish the validity of an existing norm.[110]

As Bulygin has observed, it seems clear that Kelsen has in mind validity as bind-

[107] For the sake of this hypothetical, it is assumed that the provisions are addressed squarely to the same type of case and it is not possible to read them harmoniously. They are exclusive and exhaustive of the same generic set of cases.

[108] Navarro and Rodríguez (n 8) 135.

[109] This is perhaps quite particular to Kelsen. If one takes applicability as meaning 'norms that are accepted as binding' then it becomes tautological that norms must be binding to be applicable and vice versa: ibid 132.

[110] Kelsen, *PTL* (n 9) 213.

ingness in these cases, rather than validity as membership.[111] After all, a norm does not cease being a *member* of the legal system by operation of desuetude. As I have argued above, desuetude is better captured by considering it a lapse of applicability rather than a lapse of validity. One may wonder, however, whether there is any connection between efficacy and applicability, given how Kelsen seems to think that efficacy is related to desuetude.

Bulygin has considered this question and answered it in the negative.[112] He has observed that if it were true that norms had to be efficacious to be applicable, this would put the cart before the horse. As efficacy is a measure of the degree to which norms are complied with and applied, they must *ex hypothesi* be applicable. Something must be applicable before being applied.

The better answer, on the issue of desuetude, is probably that inefficacy can lead to inapplicability. Norms that become increasingly inefficacious are more likely to fall prey to a conventional understanding that the norm has fallen into desuetude. Where this happens, the norm is not merely inefficacious, it is inapplicable. A norm being applied implies (by necessity) that it is applicable,[113] something being unapplied may imply (contingently) that it is inapplicable, depending on the scope and duration of the period of disapplication.

8.4.4. Summary and Conclusion

The formal differences between membership (D-validity; the belonging of a norm to a particular legal system), applicability and bindingness (N-validity; the capacity of a norm to make conduct obligatory) have been summarised by Bulygin as follows:

> All three concepts are relational … but the relations differ. The concept of membership designates a four-place relation between a legal system, an empowering norm of this system, an act of norm creation, and a norm created by means of this act. … The relation of bindingness is a two-place relation, defined by Kelsen as a relation

[111] Eugenio Bulygin, 'The Problem of Legal Validity in Kelsen's Pure Theory of Law' in Carlos Bernal and others (eds), *Essays in Legal Philosophy* (OUP 2015) 316.

[112] ibid 322.

[113] More technically, it is a necessary condition of application for something to be applicable.

> between a norm and the validity (bindingness) of another norm. The concept of applicability designates a five-place relation between a legal system, an application norm that belongs to this system, a court, a case, and a norm that is to be applied.[114]

Much of the confusion around validity is because of a failure to disambiguate between membership and bindingness. The term 'validity' is thus somewhat unhelpful. The three properties identified in the quote above are, however, analytically distinct. Again, of particular import in this thesis is the distinction between applicability and membership.

As regards the relationships between the concepts discussed here. It was first argued that validity and existence are not coextensive. This is partly due to a failure to distinguish between the different senses of D-validity and N-validity. It was also argued that applicability and existence are not coextensive either. Therefore, whatever the existence of legal norms is, it is not exhausted by either validity or existence. This is not to say that validity and applicability have no relationship with the existence of legal norms; I am merely claiming that whatever that relationship is, it is not one of material equivalence.

The relationship between validity and applicability is more complex. I argued above that a norm's membership (D-validity) of a legal system is often sufficient but always unnecessary for its applicability. The link between bindingness (N-validity) and applicability does not require as much detailed consideration. It was argued above that N-validity itself is dependent on efficacy and that there is no link of necessity or sufficiency between efficacy and applicability. It may therefore be said that there is *no strict equivalence relationship between applicability and validity*. This is well captured by Navarro and Rodríguez in the following quotation:

> [A] norm may belong to different systems: the applicable system for a case *c* and the legal system (ie, the set of valid norms). However, this possibility is frequently overlooked as a consequence of the confusion of applicability and membership in the legal system. If a judge imposes a certain order to solve a conflict between two norms equally applicable to a certain case, the weaker norm will be disregarded for the decision she has to make. But this does not mean that such a norm will be eliminated

[114] Bulygin, 'The Problem of Legal Validity in Kelsen's Pure Theory of Law' (n 111) 321. Bulygin references Kelsen's definition of bindingness, despite himself criticising that understanding:

> from all the systems to which it belongs. It can survive as a member in the legal system and contribute to the solution of other cases in the future. Similarly, the normal way to eliminate a norm from the legal system is through an act of derogation, but even if a norm is suppressed in this way, this does not mean that such a norm cannot survive as a member in the applicable system with respect to certain cases.[115]

An analysis of legal systems that substantiates how laws can belong to different sets in different ways is undertaken in chapter 9. For now, it suffices to note how validity and applicability relate to the three theoretical questions under consideration at present. The effects question can ultimately be seen as turning on a choice between these relations. The 'validity' answer bears on the question of membership of a legal norm within a legal system. What this means is that the norm(s) putatively created by an unconstitutional statute were never valid members of the legal system. This coheres with von Wright's idea that an invalid norm marks a recession from the normative system and effectuates no normative changes.[116] However, I also argued above that existence/validity and applicability are conceptually separable properties of legal norms. This being the case, it does not necessarily follow that these invalid norms are inapplicable.

The applicability answer to the effects question, by contrast, does not purport to make any comment on the consequences of membership or non-membership of a legal system for a legal norm. Instead, it would prevent unconstitutional norms being used as legal justifications in cases. Again, if validity and applicability are conceptually separated then this does not pose any categorisation problems. Crucially, this allows norms that are legally invalid to remain members of a set of legal norms that is of legal relevance, even if it is not the membership set of the system. If all that is happening is that precedent narrows the scope of application of the norm (even if it lowers it to nothing), this still entails no bearing on the validity of the norm and, as such, does not go to any issues surrounding the place of invalid legal norms in the legal order. Thus, it is possible to preserve legal validity while modifying the scope of applicability of a legal norm.

---

[115] Navarro and Rodríguez (n 8) 139.

[116] von Wright (n 76) 203.

8.5. THE TIMING OF UNCONSTITUTIONALITY

When does unconstitutionality start, and for what duration(s) is it effective?

The timing of unconstitutionality is important not for the theoretical reasons of the grounds and effects questions; rather, it is important because of practical considerations, such as the scope for other legal institutions to react to the unconstitutionality effectively. However, it is not entirely removed from the theoretical questions considered above. For example, the judicial intervention answer to the derivation question makes a natural pair for the 'present' answer to the origin question considered below. In the same vein, the legislative nullity view is naturally accompanied by the 'past' answer to the same question.

I divide the timing question into two sub-questions: the *origin* question, and the *duration* question. These are considered in turn below.

8.5.1. The Origin Question

At what time does an unconstitutionality begin to exist? When does something start being unconstitutional?

The temporal origin question is this: 'when does a law become unconstitutional?' Answers to this question are relatively simple, and can be divided into two camps: *past*, and *present*. As mentioned above, this is intimately connected to the derivation question: if a law is unconstitutional in virtue of a judge's saying so (the judicial intervention view) then the temporal origin can only be the present. Conversely, if a law is unconstitutional in virtue of the constitution itself, without the necessity for intervention by any legal official or institution (the legislative nullity view) then the temporal origin can only be some point in the past because this reduces the role of the judge to noticing and declaring a constitutional fault that preceded the case.

8.5.1.1. Unconstitutionality from the Future

In many respects the origin question is relatively straightforward, and so the remainder of this section will briefly explain an odd aberration whereby unconstitu-

tionality can originate from a *future* point in time. It may sound bizarre to suggest that the origin of an instance of unconstitutionality can lie in the future, rather than the past or present. After all, the past and present are susceptible to human agency in a way that the future is not. So how could law, a fundamentally artefactual concept, be manipulated in the future?

This difficulty is more common than it might seem, and it will be explored more fully in Chapter 10 where I consider the 'retrocausality problem'. For now, consider an instantiation of the legislative nullity view wherein judges interpret the constitution and apply those interpretations to decide constitutional cases, but it is still considered that the constitution (not the judge) ultimately renders the law unconstitutional. Interpretation is plausibly a creative exercise,[117] and so the judges are adding to the constitutional content through this practice. However, say this jurisdiction also maintained that the point of origin of the unconstitutionality of a piece of legislation was the date that piece of legislation was promulgated (the 'past' answer). Ordinarily, this will not pose an issue; say, in a straightforward case where a jurisdiction's constitution came into force in 1900, and a 1907 statute is found to be unconstitutional in 2007. The point of origin here must be 1907, all other things held equal.

Now say the example is complicated slightly. What if there was a major judicial reconsideration of a relevant constitutional provision in 1957? This would add new normative content to the constitution over and above the content present in 1900.[118] Assume that the 1907 statute passed muster until this reconsideration in 1957, whereupon it became unconstitutional. Its unconstitutionality should now date to either 1957; however, on the legislative nullity view, this is ignored and the point of origin is still 1907.[119] This becomes yet more anomalous still if you con-

[117] Joseph Raz, 'On the Authority and Interpretation of Constitutions: Some Preliminaries' in Larry Alexander (ed), *Constitutionalism: Philosophical Foundations* (Cambridge University Press 1998) 177.

[118] This is in effect the problem of 'creeping unconstitutionality' referenced above.

[119] It is interesting that this occurs for judicial reconsiderations of provisions, but not formal amendments. Say that instead of a judicial re-interpretative process in 1957, the Constitution was instead

sider this situation from the perspective of an official of that legal system in 1908. The 1907 statute is unconstitutional in 1908. For a legal official in 1908, the statute is therefore unconstitutional 'now', but this is completely unknowable, because judicial reconsideration of the relevant constitutional provision will not occur until 1957. In effect, this requires that communities can create laws and legal systems that entail legal facts that are, to that community, unknowable. It is quite natural to assume that this is false. The law may be complex, but all of it is in principle knowable even if it is not actually entirely known. The idea of law that is, for a particular duration, completely unknowable is thus an unintuitive consequence of the 'past' answer where it co-occurs with an even minimally creative account of judicial interpretation.

In some respects this is not just a feature of law so much as it is a feature of our way of talking about time. Terms such as 'now' or 'in the past' are relative to a temporal position; something that is 'in the past' in 2018 may not be in the past from the perspective of 1918. It seems unproblematic to say that a law is unconstitutional in the past, by operation of an adjudicative result that occurs now, merely because of the tautology that no-one's current perspective is in the past and so no-one actually experiences the anomalous 'back-to-the-future' effect described in the paragraph above. However, we only get this result by privileging the present moment in our analysis. If we shift our perspective backwards in time, it looks like the future is determining results that are obtaining in the present. As stated previously, I will examine this curious feature of unconstitutionality further in chapter 10. For now, it suffices to see how a point of origin for unconstitutionality can be either 'past' or 'present', and the difficulties that the 'past' model can encounter.[120]

formally amended in 1958 to include a provision that made our 1907 statute unconstitutional. Even on the 'past' answer, the origin here would be 1958. There are knotted corollary issues here, such as the distinction between constitutional text and constitutional content. I do not propose to explore those in detail. Suffice it to note that even a committed legislative nullity view taking the 'past' answer to the origin question would not counter-factually insist that legislation became unconstitutional before a formal amendment was passed that made it unconstitutional.

[120] This is not to say that there are not difficulties with the 'present' model either: *Linkletter* and its progeny are good examples of the difficulties with this view.

8.5.1.2. The Simulation of Retrospective Effect

The 'past' answer to unconstitutionality can only *simulate* retrospectivity. It does not provide an account of retrospective law *per se*. Recall that I defined retrospectivity as changing the legal status of actions that were undertaken *before the law came into effect*. On this view neither a judicial intervention nor a legislative nullity model requires an account of retrospectivity. This is so if either of them is paired with the 'past' answer. This is noted by Munzer, who points out with respect to the judicial intervention view that:

> An overruling decision either changes law or it does not. If it does, the theory is false and the concept of 'Blackstonian retroactivity' does not apply. If it does not, the overruling decision corrects prior erroneous interpretations of what the law is but cannot alter the legal rule retroactively; the theory offers only an account of the retroactive application of the latest judicial 'interpretations' or 'evidence' of what the law is, not of retroactive law.[121]

This observation applies equally to the legislative nullity view,[122] which holds that either a law did or did not exist for a particular duration. At no point does that theory purport to reach back into the past and change status of laws prior to the law's enactment (or lack of enactment).[123] Rather, the issue in both cases rests on the ability of laws that are invalid or inapplicable (depending on the contingent answer to the 'effects' question) to provide justification for institutional action taken in the past, not to revise the juridical status of actions that were taken before the law was taken to exist.

---

[121] Munzer, 'Retroactive Law' (n 91) 375.

[122] A more detailed discussion of this point can be found at pages 307–308.

[123] This is a point that has been indirectly acknowledged by the judiciary on occasion. For example, in considering the 'default' position regarding retrospectivity of declarations of unconstitutionality, the Constitutional Court of South Africa had the following to say: 'Because the order of the High Court declaring the section invalid as well as the confirmatory order of this Court were silent on the question of limiting the retrospective effect of the declaration, the declaration was retrospective to the moment the Constitution came into effect. That is when the inconsistency arose. As a matter of law the provision has been a nullity since that date.' *Moise v Greater Germiston Transitional Local Council* [2001] ZACC 2, (2001) 4 SA 1288 [13] (Kriegler J).

8.5.2. The Duration Question

For what span of time is unconstitutionality effective?

There are two possible answers to the duration question: *prospectively only* or *prospective and retrospective*. This question is nominally conceptually independent of the effects question, but in practice it is deeply linked to it. The origin question addresses itself to the temporal point at which unconstitutionality may be derived. It is thus a temporal foil to the derivation question. The duration question is more in the nature of a temporal foil to the effects question. It addresses itself to the span of time for which those effects will obtain.

If the effect of unconstitutionality is invalidity, and invalidity means voidness, then it seems to be the case that unconstitutionality has eternal duration; that is, it takes effect both prospectively and retrospectively. Conversely, if the effect of unconstitutionality is inapplicability, then an answer is not presumed either way. There is nothing about inapplicability to say that it cannot be imposed retrospectively, and nothing to say that it must be imposed retrospectively, either.

To see how question has significance independent of the origin question, it is useful to examine the order recently granted by the South African Constitutional Court in *Levenstein*.[124] This case concerned the unconstitutionality of section 18 of the Criminal Procedure Act 1977. The reasons for the unconstitutionality are not of particular concern here; what is important is that one of the questions the court had to answer was what period the unconstitutionality would cover. There were two options presented to the Court by counsel: either the unconstitutionality should date from the effective date of the Final Constitution (4 February 1997), or the effective date of the Interim Constitution (27 April 1994).[125] The reason for this is that section 18 had been replaced by section 27(1) the Criminal Law Amendment Act 1997, but that replacement provision had been deemed effective from the 27 April 1994.

[124] *Levenstein v Estate of the Late Sidney Lewis Frankel* [2018] ZACC 16.

[125] ibid [30], [77].

From an origin perspective, this is difficult to parse, though the Court couches it in those terms.[126] Consider that the Court was exercising a review power granted to it under the 1997 Constitution. The earliest possible date for an origin point for the unconstitutionality is, therefore, 1997. Whether the unconstitutionality has retrospective effect *beyond* that origin point is therefore a question of temporal duration. In the end, the Constitutional Court found that the unconstitutionality 'dated from' 1994.[127] Strictly speaking, this cannot be the case. The court was exercising a power to determine unconstitutionality that, itself, dated from 1997. Using this power to say that an instance of unconstitutionality *originated* in 1994 seems incorrect. It is not, however, impossible to say that the exercise of a 1997 power can make it the case that a law in 1994 can be deemed unconstitutional by a power that only came into existence in 1997. This kind of retrospectivity is not revolutionary in law. At any rate, it is this distinction that the origin/duration temporal questions are attempting to capture.

8.6. COMBINING ANSWERS TO THE THREE QUESTIONS

In this section, I will present theoretically possible models of unconstitutionality based on the possible combinations of answers to the questions above. For this purpose, the temporal question is split into the origin and duration questions. This entails four questions total, each of which has two possible answers. Mathematically, this entails 16 possible options.[128] These are presented in table 8.1.

[126] 'The next issue for consideration is the date from which the declaration of invalidity should run.' *Levenstein v Estate of the Late Sidney Lewis Frankel* (n 124) [77].

[127] ibid [89].

[128] This is given by: $\binom{2}{1} \times \binom{2}{1} \times \binom{2}{1} \times \binom{2}{1} = 2 \times 2 \times 2 \times 2 = 16$.

Grounds	Effects	Origin	Duration	Instantiating Cases and Doctrines
Legislative Nullity	Invalidity	Past	Retrospective & Prospective	Ireland (*Murphy*); South Africa (*Ferreira*); Canada (*Big M*); United States (*Marbury*, *Norton*); India (*Sagir Ahmad*, *Vij*).
Legislative Nullity	Invalidity	Past	Prospective	Ireland (*A*, *Damache*).
~~Legislative Nullity~~	~~Invalidity~~	~~Present~~	~~Retrospective & Prospective~~	- -
~~Legislative Nullity~~	~~Invalidity~~	~~Present~~	~~Prospective~~	- -
Legislative Nullity	Inapplicability	Past	Retrospective & Prospective	Canada (suspended declaration); South Africa (suspended declaration); Ireland (pre-constitution law); India (*Ambica Mills*); India (doctrine of eclipse/pre-constitution law).
Legislative Nullity	Inapplicability	Past	Prospective	- -
~~Legislative Nullity~~	~~Inapplicability~~	~~Present~~	~~Retrospective & Prospective~~	- -
~~Legislative Nullity~~	~~Inapplicability~~	~~Present~~	~~Prospective~~	- -
Judicial Intervention	Invalidity	Past	Retrospective & Prospective	- -
Judicial Intervention	Invalidity	Past	Prospective	- -
Judicial Intervention	Invalidity	Present	Retrospective & Prospective	- -
Judicial Intervention	Invalidity	Present	Prospective	- -
Judicial Intervention	Inapplicability	Past	Retrospective & Prospective	- -
Judicial Intervention	Inapplicability	Past	Prospective	- -
Judicial Intervention	Inapplicability	Present	Retrospective & Prospective	United States (*Mellon*, as-applied challenges).
Judicial Intervention	Inapplicability	Present	Prospective	United States (*Linkletter*).

Table 8.1: Combinations of Answers to Questions on Unconstitutionality (Strikethrough Indicates Views Containing Tensions)

Of course, this is just an array produced mathematically at this point. Constraints of theoretical consistency make some of these options difficult to work. One such constraint already mentioned above is the interconnectedness between the derivation and origin questions. It must be stressed that a violation of this constraint is not an outright contradiction, but it is a tension. Legislative nullity fits best with a view that the law is invalid from some point in the past because the constitution is taken to operate on legislative provisions when they are enacted. In a jurisdiction that adheres to the idea of legislative nullity, the judiciary will likely present themselves as discovering, rather than creating, constitutional norms. This means that the constitution has all its normative content from its inception, except where formal amendments add or delete additional textual content. Thus, when determining the point at which a piece of legislation came into contradiction with the constitution, the date must be the enactment of that legislation as the constitutional norm, *ex hypothesi*, existed before the legislation was passed.

The practices of the various jurisdictions from Part II are arranged below under three broad thematic headings: 'voidness *ab initio*', 'applicability-regulating mechanisms' and 'prospective-only mechanisms'. I place an abbreviated form of the answers to the theoretical questions in this chapter beside each model (the designation 'XX' means that the theme does not necessarily require an answer to that question). The abbreviations used are: 'Ln' (Legislative Nullity), 'Ji' (Judicial Intervention), 'Va' ((In)validity), 'Ap' ((In)applicability), 'Pa' (Past), 'Pr' (Present), 'Rt' (Retrospective), and 'Pr' (Prospective).[129]

An additional point to note is that I have described suspended declaration practices as going to applicability. This may seem surprising, given how courts often couch these declarations in expressly temporal terms. However, this temporal vocabulary is, itself, confused. The two jurisdictions that have firmly established the practice of suspending declarations (Canada and South Africa) also have firmly estab-

[129] 'Pr' is used as an abbreviation twice. It means 'present' if it is the third character, and 'prospective' if it is the fourth or fifth.

lished constitutional supremacy clauses,[130] which the courts have acknowledged as effectively regulating validity autonomously.[131] It is a contradiction to assert both that a court has no role in constituting validity and that a court can determine when validity is constituted.

As I argued above, the validity and applicability of a legal norm can be treated separately in some cases. This distinction can be used to rationalise the practice of suspended declarations against the background of such supremacy clauses. The courts often cast suspension in temporal terms because they lack the conceptual vocabulary to draw a distinction between validity and applicability. The analysis in this chapter attempts to supply this conceptual vocabulary, and this vocabulary can be used to defuse the contradiction identified above. There is nothing difficult in simultaneously asserting that the constitution regulates validity, but that the courts may treat invalid norms as applicable in certain cases. This is effectively what a suspended declaration achieves, and so practices of suspension are counted as applicability-regulating mechanisms.

8.6.1. Voidness *Ab Initio* (Ln | Va | Pa | Rt & Pr)

This is the most widespread model of unconstitutionality. It has occurred, at the very least rhetorically, in every jurisdiction in this survey. The harshness that can be produced by the application of this view has frequently led courts to develop doctrines avoiding or curtailing unconstitutionality, or limiting the breadth of effect of unconstitutionality once apprehended.

The approach of the Irish Supreme Court in *Murphy* is best understood through this lens. Recall from chapter 2 that Article 15 of the Irish Constitution supports

---

[130] Section 2 of the Constitution of South Africa: 'This Constitution is the supreme law of the Republic; law or conduct inconsistent with it is invalid, and the obligations imposed by it must be fulfilled.'
Section 52 of the Constitution Act 1982: 'The Constitution of Canada is the supreme law of Canada, and any law that is inconsistent with the provisions of the Constitution is, to the extent of the inconsistency, of no force or effect.'

[131] *R v Big M Drug Mart Ltd* [1985] 1 SCR 295 (SCC); *Ferreira v Levin* [1995] ZACC 13, (1996) 1 SA 984; *Geldenhuys v State* [2008] ZASCA 47, (2009) 1 LRC 294.

both the legislative nullity and invalidity views. In *Murphy* the Irish Supreme Court further made clear that the answer to the 'origin' question was the point in the past where the unconstitutional legislation had putatively been enacted. Henchy J conveyed this through the memorable analogy of a 'judicial death certificate' dated to 1937 (the year the Irish Constitution came into force) in the case of pre-constitution law and the date of enactment of the legislation in the case of post-constitution law.[132]

This is also the theory that best describes the South African doctrine of 'objective invalidity'. Recall that in *Ferreira v Levin* the Constitutional Court observed that a declaration of unconstitutionality 'does not invalidate the law; it merely declares it to be invalid'[133] and that this validity was to be construed objectively. Although in chapter 7 it was noted that more recent decisions on the doctrine are showing some conceptual discomfort with its strictures[134] it nevertheless remains the law and thus the default position in South Africa.[135]

This theory is also supported by the early statements in the US Supreme Court in cases like *Marbury v Madison*[136] and *Norton v Shelby County*.[137] Strong statements on invalidity and voidness in those cases mean that the early answer of the US Supreme Court to the 'effects' question was *invalidity*. Furthermore, support for the Blackstonian view of law in cases such as *Erie*[138] are good evidence that the early

---

[132] Only the point regarding pre-constitution law was made directly by the judge, but no reason was given why the 'judicial death certificate' imagery would not apply equally to post-constitution law. *Murphy v Attorney General* [1982] IR 241 (SC) 307. McGuinness J seemed to apply the imagery in this more expansive way subsequently, noting its ubiquity: *A v Governor of Arbour Hill Prison* [2006] IESC 45, [2006] 4 IR 88 [189].

[133] *Ferreira v Levin* [1995] ZACC 13, (1996) 1 SA 984 [27].

[134] *Department of Transport v Tasima (Pty) Ltd* [2016] ZACC 39, (2017) 1 BCLR 1.

[135] 'The general assumption . . . is that an unconstitutional provision is invalid with immediate effect and that a party wishing the Court to suspend its order of invalidity must provide persuasive reasons for the Court to do so.' Stuart Woolman and Michael Bishop (eds), *Constitutional Law of South Africa* (2nd edn (rev), Juta 2013) 12–78.

[136] *Marbury v Madison* 5 US 137 (1803).

[137] *Norton v Shelby County* 118 US 425 (1886).

[138] *Erie R Co v Tompkins* 304 US 64 (1938).

view of the US courts on the 'derivation' question was legislative nullity. The modern view is likely better stated in the *Mellon*[139] case, as described in chapter 2.

This theory is also supported by some cases in the Supreme Court of India, such as *Sagir Ahmad*[140] and *Vij*.[141] In those cases the Court affirmed what it has characterised as the 'American' view that unconstitutional legislation is void *ab initio*. Notably, these cases took place under Article 13(2) of the Constitution of India, and the treatment of pre-constitution law under Article 13(1) differs. This distinction in treatment is parsed through the analytic device of the 'doctrine of eclipse',[142] which notionally only holds for pre-constitution law.[143] There is some uncertainty, however, over whether the doctrine extends to post-constitution law. This reflects a more general lack of certainty as to the theoretical underpinnings of unconstitutionality in the Indian legal system.

Finally, it seems likely that this theory describes the approach of the Supreme Court of Canada in *Big M Drug Mart Ltd*. Recall that in that case the Court found that:

> Section 52 sets out the fundamental principle of constitutional law that the Constitution is supreme. The undoubted corollary to be drawn from this principle is that no one can be convicted of an offence under an unconstitutional law. … Any accused, whether corporate or individual, may defend a criminal charge by arguing that the law under which the charge is brought is constitutionally invalid.[144]

As I argued in chapter 4, the way in which the court envisages standing to be irrelevant here suggests that the constitution regulates its own self-executing standards of unconstitutionality, and the judiciary just act as ciphers for this process.[145]

It is more difficult to classify with confidence the Canadian position on the ef-

---

[139] *Massachusetts v Mellon* 262 US 447 (1923).

[140] *Saghir Ahmad v State of UP* [1954] INSC 89, [1955] SCR 707.

[141] *Rakesh Vij v Raminder Pal Singh Sethi* [2005] INSC 522.

[142] *Bhikaji Narain Dhakras v State of Madhya Pradesh* [1955] INSC 48, [1955] 2 SCR 589.

[143] *KK Poonacha v State of Karnataka* [2010] INSC 706.

[144] *R v Big M Drug Mart Ltd* [1985] 1 SCR 295 (SCC) [38].

[145] Grellette (n 21) 27.

fects question. The answer with respect to the derivation of unconstitutionality in Canada is clear: it is undoubtedly legislative nullity. What the Constitution *does* is less clear. The phrase 'of no force or effect' suggests something more like inapplicability than invalidity. However, the Supreme Court of Canada has consistently rejected the view that it allows for a practice of legislation being held 'unconstitutional as applied'.[146] Additionally, the fact that Big M did not need a substantive right to freedom of religion also may suggest that applicability is not relevant to the Canadian analysis. It was sufficient for Big M to argue *any* reason (whether applicable to its own circumstances or not) for the legislation to be invalid and, from there, it merely needed to make the simple step of objecting to prosecution under an unconstitutional law. The fact that the reason for the unconstitutionality was not itself applicable to Big M's circumstances suggests that the Canadian courts did not in that case consider that unconstitutionality regulates applicability.

8.6.2. Applicability-Regulating Mechanisms (XX | Ap | XX | Rt & Pr)

What each of these doctrines have in common is that, for one reason or another, a validity analysis is irrelevant. This may be because validity is determined by reference to a now-extinct or foreign legislature and legal system (pre-constitution law) or because a constitutional supremacy clause has already settled it and the court has no power to vary that finding (suspended declarations in Canada and South Africa). It may also simply be a feature of a judiciary notionally not having the power to regulate validity at all (as-applied challenges in the United States).

8.6.2.1. Pre-Constitution Law in Ireland and India (Ln | Ap | Pa | Rt & Pr)

This is the theory that best captures the approach in both Ireland and India to constitutional review of laws that pre-date the enactment of the constitutional text. The Irish courts seem to prefer the view that laws that were inconsistent with the 1937

[146] *R v Nur* 2015 SCC 15 (CanLII), [2015] 1 SCR 773 [72]–[73]; *R v Ferguson* 2008 SCC 6 (CanLII), [2008] 1 SCR 96 [72]; *R v DeSousa* 1992 CanLII 80 (SCC), [1992] 2 SCR 944, 955.

Constitution, but which were enacted before its promulgation, could be deemed valid or invalid only by reference to the Parliament that enacted them (either the Oireachtas of the Irish Free State, or the Parliament in Westminster).[147] As a result, the Irish courts seem to consider whether pre-constitution legislation continues to be *applicable*, rather than whether it is valid.[148] In India, the doctrine of eclipse[149] seems to provide good evidence that it is the applicability, rather than the validity, of pre-constitution law that is at issue. This is because it allows for unconstitutional pre-constitution law to regain its applicability where this impediment to its constitutionality is subsequently removed or obviated.

It is worth noting that this treatment of pre-constitution law suggests a theoretical view known as 'exclusive legal positivism' (ELP). This is often contrasted with 'inclusive legal positivism' (ILP).[150] Two of the central claims of modern legal positivism are: (1) the claim that law is a social fact, and (2) that legal validity is a matter of the source of a legal norm within the legal system, not its merits. ELP maintains this quite strictly. ILP allows that the merits of a legal norm may be incorporated by reference to the source-based analysis; in other words, where the source of the legal norm incorporates a merit-based standard, then those merits become legally relevant.

Regarding unconstitutionality, ILP seems to conceptually require the conclusion that unconstitutional norms are invalid and inapplicable *ab initio* because the legal norms at play really *did* conflict *ab initio*; whatever moral standards the relevant constitutional provision incorporated by reference always required that the statutory norm was invalid, because most people think that moral principles do not

---

[147] *Geoghegan v Institute of Chartered Accountants in Ireland* [1995] 3 IR 86 (HC).

[148] *The State (Sheerin) v Kennedy* [1966] IR 379 (SC); *The People (DPP) v MS* [2003] 1 IR 606 (SC); *JP v Director of Public Prosecutions* [2009] 3 IR 215 (HC).

[149] *Bhikaji Narain Dhakras v State of Madhya Pradesh* (n 142).

[150] Michael Guidice, 'Unconstitutionality, Invalidity, and Charter Challenges' (2002) 15 Can J L & Jurisprudence 69; Waluchow, 'Four Concepts of Legal Validity' (n 75) 134; Wil Waluchow, *Inclusive Legal Positivism* (Clarendon Press 1994).

change over time.[151] Section 52 of the Canadian Constitution Act 1982 is a good example of this.[152]

Ireland and India seem better captured by ELP. The reluctance of both Supreme Courts in those jurisdictions to adjudicate on the *validity* of pre-Constitution law reflects that such law has its *original*[153] source of validity in another legal system (the Irish Free State, Dominion of India, or the United Kingdom). The court cannot make a claim on the sources of validity in those other legal systems, and therefore it does not adjudicate on the validity of pre-Constitution law. Post-Constitution law, however, is subject to a validity analysis as it has its source in the same legal system as the adjudicating court.

I believe ELP has the better of this argument. The difficulty with ELP, however, is in accounting for the *voidness* of invalid law as opposed to its sheer invalidity. Because ELP accepts that judges *create* more constitutional law when interpreting moral criteria incorporated in the Constitution,[154] by its own lights it should require that a law be invalidated from the date of the creation of the new norm that made it invalid. This is not what is reflected in legal practice under the legislative nullity view, which continues to maintain that unconstitutional statutes are void, and that voidness is *ab initio*.

---

[151] Michael Guidice, 'Unconstitutionality, Invalidity, and Charter Challenges' (2002) 15 Can J L & Jurisprudence 69, 81; Grellette (n 21) 27.

[152] The doctrine of objective invalidity in South Africa might also support the view that South Africa is also best explained by ILP.

[153] Of course, these laws also have a source in the legal system of the Republic of Ireland or Republic of India, respectively. It is just not their ultimate source. It would therefore be open to the Irish and Indian courts to declare these norms invalid-in-Ireland or invalid-in-India. This would be consistent with ELP, and is one interpretation of what those courts are doing when ruling on the constitutionality of pre-Constitution law, but it nevertheless seems significant that they have avoided the terminology of 'validity' and have in some cases applied different standards and rules to pre-Constitution law.

[154] Grellette (n 21) 29–30; Joseph Raz, *Ethics in the Public Domain* (rev edn, OUP 1995) 242.

8.6.2.2. Ambica Mills in India (Ln | Ap | Pa | Rt & Pr)

This is the more general theory in India, after the *Ambica Mills*[155] case brought significant clarity (if not consistency with the constitutional text) to Indian practice. In that case, Mathew J endorsed voidness as a relative rather than an absolute concept:

> [V]oidness is not *in rem* but to the extent only of inconsistency or contravention, as the case may be of the rights conferred under Part III. Therefore, when article 13(2) uses the expression 'void', it can only mean, void as against persons whose fundamental rights are taken away or abridged by a law. The law might be 'still-born' so far as the persons, entities or denominations whose fundamental rights are taken away or abridged, but there is no reason why the law should be void or 'still-born' as against those who have no fundamental rights.[156]

Although there are semantic oddities in describing, by implication, voidness as being *in personam*, the general thrust of this judgment is that unconstitutional law may continue to be applicable to certain cases. In particular, in the context of the Indian constitution, it may be applicable against non-citizens, who cannot avail of the fundamental rights protections of Part III of that constitution.

8.6.2.3. Suspended Declarations of Unconstitutionality (Ln | Ap | Pa | Rt & Pr)

As argued above, it seems more plausible to consider the suspended declaration of unconstitutionality not as holding the effect of invalidity at bay until a point in the future, but rather maintaining the applicability of unconstitutional law for the present. This seems to be the best way to harmonise this practice with the existence of constitutional supremacy clauses that purport to regulate validity *ex proprio vigore*. Such clauses are not compatible with temporal modifications to the effect of unconstitutionality. This being the case, and notwithstanding the language judges themselves employ, suspended declarations are not a modification of the temporal question as I have styled it in this chapter. This reading also allows suspension to co-exist somewhat more consistently with the voidness *ab initio* practices in each of the jurisdictions styled above.

---

[155] *State of Gujarat v Shri Ambica Mills* [1974] INSC 72, [1974] 3 SCR 760.

[156] ibid 777–78.

The merits and demerits of the suspended declaration, and the clarity than an applicability analysis can bring to the development of this remedy, are more fully considered in chapters 10 and 11.

8.6.2.4. As-Applied Challenges (Ji | Ap | Pr | Rt & Pr)

This is the theory reflected in the US practice of holding provisions of legislation unconstitutional only as they apply to particular cases, rather than deeming them facially invalid. This distinction is grounded in the *Salerno*[157] case, discussed in chapter 3. As was discussed in that chapter, the ambit of *Salerno*, and the facial/as-applied distinction in general, remains somewhat unclear. While the Supreme Court often praises the as-applied challenge rhetorically, the results it achieves are more consistent with its statements on facial challenges.

The facial/as-applied distinction is not the only piece of data that suggests the United States federal courts favour applicability analysis. In chapters 2 and 3 I pointed to several cases, such as *Massachusetts v Mellon*,[158] in which the federal courts declared that they have no power to regulate the validity of unconstitutional statutes and that they can only enjoin their enforcement and actions taken pursuant to them. This requires that the federal courts have a power to modify the applicability of legal norms.

The United States is the only jurisdiction that has at any point opted for a theory that did not take legislative nullity as an answer to the derivation question. This is perhaps reflective of eagerness on the part of common law judges to distance themselves from politically volatile conclusions, as can occasionally arise from the application of unconstitutionality. It may be more attractive to present unconstitutionality as a *fiat*, which binds a regretful judge, in these circumstances. Additionally, it may also suggest that the judicial intervention approach will be more attractive, or at least more likely to develop, where there is no clear Constitutional statement allow-

[157] *United States v Salerno* 481 US 739 (1987).

[158] *Massachusetts v Mellon* (n 139).

ing courts to find laws unconstitutional.

8.6.3. Prospective-Only Mechanisms (XX | XX | XX | Pr)

8.6.3.1. The Legislative Nullity Variant: Ireland (Ln | Va | Pa | Pr)

This is the theory endorsed currently by the Irish courts in *A v Governor of Arbour Hill Prison*[159] and *Damache v Director of Public Prosecutions*.[160] The general rule in these cases differs from the original rule endorsed in *Murphy*[161] insofar as it makes retrospective effect the exception rather than the norm. The case law following *Damache* has established that, at least in criminal proceedings, an accused must satisfy three criteria to benefit retrospectively from a declaration of unconstitutionality. They must: (1) have raised the issue themselves at trial, and (2) not have taken steps that could be taken as acquiescing to the charge, such as pleading guilty. Additionally, (3) the proceedings must be non-final. This makes the default stance that the unconstitutionality will apply prospectively only, except where these relatively strict conditions are met.

8.6.3.2. The Judicial Intervention Variant: United States (Ji | Ap | Pr | Pr)

This was the theory endorsed by the US Supreme Court in the *Linkletter* era. The key differentiating feature of this jurisprudence was its *prospective* effect. The classification of this practice as one modifying applicability must be somewhat tentative, as the United States courts are occasionally rhetorically inconsistent on that point. In any case, it cannot be doubted that this case law endorsed a prospective-only standard for unconstitutionality. This standard was, however, ultimately abandoned by

[159] *A v Governor of Arbour Hill Prison* [2006] IESC 45, [2006] 4 IR 88.

[160] *Damache v Director of Public Prosecutions* [2012] IESC 11, [2012] 2 IR 266.

[161] *Murphy v Attorney General* [1982] IR 241 (SC).

the Court in criminal,[162] civil,[163] and *habeas corpus*[164] proceedings.

8.7. CONCLUSION

To recapitulate, the framework of unconstitutionality developed in this thesis centres on three questions:

(1) The derivation question: *Whence does unconstitutionality derive? In virtue of what power, or failure of the exercise of power, is unconstitutionality established?*

(2) The effects question: *What does unconstitutionality do? In what way does unconstitutionality modify legal norms?*

(3) The temporal question, which breaks down into two further sub-questions:

   (a) The origin question: *At what time does an unconstitutionality begin to exist?*

   (b) The duration question: *For what span of time is an unconstitutionality effective?*

This framework was shown to generate sixteen different answers, which can then be arranged in a matrix to synthesise models of unconstitutionality. These can then be mapped back on to the jurisdictions to rationalise legal practice and isolate points of distinction and transformation both intra- and inter-jurisdictionally.

This analysis identified several trends among the jurisdictions. The most significant of which is the treatment of unconstitutional law as being void *ab initio* at some point in time in every jurisdiction studied in this thesis. This view is harsh in at least two dimensions: (1) its effect is to destroy the law wholesale (finding it

[162] *Griffith v Kentucky* 479 US 314 (1987).

[163] *Harper v Virginia Department of Taxation* 509 US 86 (1993).

[164] *Teague v Lane* 489 US 288 (1989).

'invalid', 'void') and (2) it does this with assumed full retrospective effect. The difficulties to which both of these dimensions can give rise are evident in virtually every jurisdiction studied in this thesis.

Another significant trend that was identified was the introduction of applicability-regulating measures in almost every jurisdiction (the possible exception here being Ireland, though that may be changing with the gradual embrace of suspended declarations of unconstitutionality). These measures allow a court to attenuate the severity of unconstitutionality and tailor its relief more finely. However, courts and commentators lack the theoretical and conceptual tools to describe these distinctions clearly, and so often they are treated as measures that modify the temporal aspect of unconstitutionality, rather than the effect of unconstitutionality itself (the suspended declaration is the prime example here). This thesis thus offers a sharper analytic lens for debating and refining these innovations, which theme is revisited in chapter 11.

In subsequent chapters, I will further develop a view of legal systems that reinforces the view that invalidity is neither ineluctable as a consequence of unconstitutionality, nor is it as severe as it is often taken to be. I will also subject the void *ab initio* view to specific criticism. This discussion will aim to build on the foundation in this chapter, establishing that invalidity and voidness are not sacred cows of unconstitutionality and that there are viable, and preferable, theoretical alternatives available.

9 The Place of Unconstitutional Law in the Legal System

CHAPTER OVERVIEW

9.1. Introduction 255

9.2. Law and Legal System 257

9.3. Legal Systems as Sets 259

9.3.1. Relating Laws to Other Laws 260

9.3.2. Relating Systems to Other Systems 262

9.4. Legal System and Legal Order 263

9.4.1. Membership vs Inclusion 266

9.4.2. Ordered and Unordered Sets 271

9.4.3. The 'Internal and External' Time of Legal Norms 273

9.5. The Consistency and Completeness of Legal Systems 274

9.5.1. Raz's Error 274

9.5.2. Bulygin on Consistency and Completeness 276

9.6. Conclusion 278

9.1. INTRODUCTION

The goal of this chapter is to situate unconstitutional law in a theoretical account of laws and legal systems. To an extent, the analysis here will depend on the 'effects' question asked in chapter 8. The property of law that is modified by unconstitutionality—its validity or applicability, or both, as the case may be—is an important piece of this puzzle. The answer to the effects question will tell us how unconstitutionality

operates on individual laws. The question of this chapter is somewhat broader and more abstract; it seeks to situate an account of the disruption achieved by unconstitutionality in a more theoretical account of legal systems. There are a few good reasons to do this:

(1) As was seen in Part II, unconstitutionality is often bound up with an analysis of validity. In chapter 8 I argued that validity should be narrowly construed as 'system membership'. This definition, through its reliance on systematicity, naturally draws attention to an analysis of legal systems being important to understand validity and, by extension, unconstitutionality.

(2) The problem of how laws persist from one time to another requires an understanding of both laws and legal systems. Persistence refers to the continuation of law through time. As most objects exist in time, they go through property changes (think of a lit candle melting) without losing their *identity* (the half-melted candle is still the same candle, even though it has undergone substantial physical change). Some property changes will destroy the object, however (an entirely liquefied candle is more a 'puddle of wax' than it is a 'candle'). Unconstitutionality entails a property change that may or may not destroy a legal norm and so is, in part, a question on the persistence of legal norms.

(3) Finally, a more thorough account of legal systems deepens understanding of the distinction drawn between validity and applicability in chapter 8 by demonstrating several ways in which laws might relate to legal systems.

Dividing the properties of legal norms in this way further entails that a picture of legal systems being made up solely of 'valid' norms is insufficient. There exist different sets within legal systems to which legal norms may belong. The set of valid norms is only *one* of these sets. This requires us to rethink legal systems in a way that is somewhat more technical than the relatively simple model that practitioners and officials within a system may tacitly accept in rationalising their own practices. Although this theory is quite conceptually involved, it is crucial to understand it

to parse unconstitutionality properly. At its most abstract, unconstitutionality is a question of how legal norms that are flawed in a certain way relate to legal systems. We cannot hope to account for this without *some* understanding of legal systems and how they can relate to legal norms generically, let alone flawed ones.

9.2. LAW AND LEGAL SYSTEM

The relationship between law and legal system is important in legal philosophy generally. Of particular relevance to this thesis is an analysis of legal validity as membership of a legal system. Given that this definition requires that one can identify legal systems otherwise than by simply classifying them as systems of valid laws, this suggests that an analysis of legal systems must be conceptually prior to an analysis of laws.[1] In his study of legal systems, Raz draws attention to precisely this point, going so far as to suggest that 'all the existing theories of legal system are unsuccessful in part because they fail to realise this fact'.[2] Gardner also suggests that Austin and Bentham's failure to answer the question 'what is law?' was a result of their giving priority to individual laws, from which systems are then derived.[3] Gardner considers that Hart rectified this error, attributing to Hart the view that 'One cannot begin by distinguishing laws and then distinguish legal systems afterwards as systems containing laws. One needs to distinguish legal systems in order to distinguish laws'.[4] Indeed, Gardner goes so far as to say that 'legal systems are the basic units of law, and laws are essential (but not the only) sub-units'.[5]

The structural features of legal systems, however, were not the focus of Hart's legal philosophy. Though he laid the groundwork for it, the detailed scrutiny of in-

---

[1] For support of this view that legal systems and laws should be subjects of separate ontological studies, see: G MacCormack, '"Law" and "Legal System"' (1979) 42 MLR 285.

[2] Joseph Raz, *The Concept of a Legal System: An Introduction to the Theory of Legal System* (Clarendon Press 1980) 2.

[3] John Gardner, 'The Legality of Law' in John Gardner, *Law as a Leap of Faith* (OUP 2012) 178.

[4] ibid 181.

[5] ibid 181.

stitutions and systems only followed his work.[6] The most detailed analyses of legal systems undertaken by more contemporary legal philosophers are those found in Raz[7] and Bulygin.[8] Both models either expressly conceive of legal systems as sets,[9] or follow a quasi set-theoretical model.[10]

In the sections that follow, I identify some structural features of legal systems on this set-theoretical understanding. In particular, I analyse the relationship of instantaneous legal systems to durative ones;[11] that is, the relationship between legal systems that last for an instant, or very short duration, and those that may last for years, decades, or even centuries.

I also note the views of both Raz and Bulygin with respect to the following two properties of legal systems:

> *Completeness:* If there is no legal question that does not have a legal solution, then the legal system is complete.
>
> *Consistency:* If no two norms that belong to the same system contradict each other, the legal system is consistent.[12]

These properties are important for this thesis because they indirectly influence whether there really is an antecedent law providing a legal answer in cases of unconstitutionality. A model of unconstitutionality endorsing the existence of antecedent,

---

[6] Seán Coyle, 'Hart, Raz, and the Concept of a Legal System' (2002) 21 Law & Philosophy 275, 283–84.

[7] Raz, *The Concept of a Legal System* (n 2).

[8] Carlos Alchourrón and Eugenio Bulygin, *Normative Systems* (Springer-Verlag 1971); see also generally: Carlos Bernal and others (eds), *Essays in Legal Philosophy* (OUP 2015).

[9] Eugenio Bulygin, 'Time and Validity' in Carlos Bernal and others (eds), *Essays in Legal Philosophy* (OUP 2015) 172–73.

[10] Coyle (n 6) 294; Raz, *The Concept of a Legal System* (n 2).

[11] Raz and Bulygin use different terminology to capture this distinction, so I use this neutral phrasing.

[12] The formulation for both these theses is taken from: Eugenio Bulygin, 'Kelsen on the Completeness and Consistency of Law' in Carlos Bernal and others (eds), *Essays in Legal Philosophy* (OUP 2015) 337.

true answers to legal questions requires that the legal system be *complete*. As I shall argue below, this is a lofty and likely unattainable commitment.

9.3. LEGAL SYSTEMS AS SETS

The questions that a theory of legal systems must answer have usefully been captured by Raz in the introduction to his own theory on the subject. These questions are: *(i)* what are the criteria for the existence of a legal system? *(ii)* what are the criteria that determine the system to which a given law belongs? *(iii)* is there a structure common to all legal systems, or to certain types of legal system?[13]

The detailed response that Raz gives to these questions is not something I will address in full here. I would, however, draw attention to one important proposition that emerges from his study in response to questions *(i)* and *(iii)*: a legal system, in Raz's view, is a *set of laws*. It is important to clarify that Raz does not himself use the word 'set' in *The Concept of a Legal System*. However, he did go on a year later to issue a statement that looks conspicuously like a concession that an analysis of legal systems must fall back on set theory:

> The problem of identity of legal systems is the quest for a criterion or set of criteria that provides a method for determining whether any set of normative statements is, if true, a complete description of a legal system. . . . A statement is a normative statement if and only if the existence of a norm is a necessary condition for its truth. A normative statement is pure if and only if the existence of certain norms is sufficient for its truth. *The set of all the pure statements referring to one legal system is called the 'total set' of that system, and every set of pure statements that is logically equivalent to the total set of a system is a complete description of that system.*[14]

This is significant, as if a legal system is a set in the *mathematical* sense of that term, then this has important consequences for the identity criteria applicable to legal systems; specifically, if they are sets, then legal systems will, as any other set, be defined extensionally by reference to the full list of their elements. Consider two

[13] Raz, *The Concept of a Legal System* (n 2) 1–2. Raz also mentions a fourth question regarding whether there are certain types of law or legal norm that recur in all legal systems but notes that this question is not essential to his study and nor has it been the focus of much jurisprudential analysis.

[14] Joseph Raz, 'The Identity of Legal Systems' (1971) 59 Ca L Rev 795, 797. Emphasis added.

ordinary mathematical sets: $A = \{1, 2, 3\}$ and $B = \{1, 2, 4\}$. These sets are clearly non-identical: $A \neq B$. This is so even if they share substantial elements; the numbers 1 and 2 in this case.

This extensional method of definition cuts the same way for legal systems, conceived of as sets. Let us call these legal system sets LS. These sets will have members; they are not empty sets. Let us say for the sake of simplicity that the elements of these sets will be laws (l).[15] Let us take two hypothetical sets: $LS_1 = \{l_1, l_2, l_3\}$ and $LS_2 = \{l_1, l_2, l_4\}$. By application of our set-theoretic definitions, these must be different legal systems: $LS_1 \neq LS_2$. This would still be so even if l_4 was only a minor restatement of l_3, in the same way that the difference between the numbers 3 and 4 mattered so much to our ordinary mathematical sets above.[16]

9.3.1. Relating Laws to Other Laws

It is well and good to say that legal systems are sets, but this does not get us very far on its own. After all, sets can merely be arbitrary lists of otherwise unconnected elements. So, one could make a set of all the legal provisions in the world that are given by section 28 of any statute. This would be a very large set, and its members would only be connected by the non-substantive feature of appearing as the 28th section in a statute. Based on the discussion thus far, this would qualify as a 'legal system' because it is a set of laws. This is patently insufficient. We thus require something more refined than simply saying legal systems are sets of laws. We need to determine the ways in which laws relate to one another within systems that interconnect them in a particular way that marks them out as legal systems. The collections of laws we call legal systems are related in this distinctive way.

---

[15] At a more granular level, legal systems contain both norms and normative statements, if not other things besides. It would unnecessarily complicate the notation used at present to account for this detail. This said, another of the defining contributions of Raz's work was to clarify the point that not all laws are norms and, thus, that an analysis of legal system taking the broader class of 'laws' as its fundamental unit was required.

[16] 'Whenever a new norm becomes a member of a legal system, we obtain a different system because the set of norms that are members of it is different'. Bulygin, 'Time and Validity' (n 9) 173.

Raz suggests that laws are joined by two relations within the legal system: a 'genetic relation' and an 'operative relation'.[17] The genetic structure of a legal system accounts for the relation of a law to another law that authorises its existence, or the repeal or modification of laws. The genetic structure provides for legal validity, as it shows how the relations between some laws will enable or disable the legal validity of other laws. The operative structure, by comparison, is concerned with the normative effects[18] produced by a law at a given moment and not the method of creation, amendment, or destruction of a law. Given that effects relate to normative content in this way, the extent to which the legal system shapes the duties and obligations of its subjects is exhausted by the operative structure. The genetic structure, however, requires more than legal norms establishing these first-order duties and obligations.[19]

The genetic structure is of importance in distinguishing random or disorganised collections of legal norms from genuine legal systems. Because the genetic structure of a legal system authorises the existence, repeal, and modification of laws, all the laws of that system will be traceable back to these change-managing laws. These are what Hart would have termed 'secondary rules'.[20] Even if these laws are themselves changed, this would have to be done in a way consistent with those laws (prior to the changes). To do otherwise would be a 'legal revolution' and would result in an

[17] Raz, *The Concept of a Legal System* (n 2) 183–85.

[18] Raz does not explain what he means by 'effect', short of suggesting that the operative structure of a legal system is 'based on its punitive and regulative relations'. ibid 185. Given this, it seems likely that he means the normative claims endorsed by the law, rather than whether those laws are in fact enforced.

[19] Bulygin's theory on legal systems also supports the view that there is a genetic and an operative structure in a legal system. Bulygin and Raz diverge, however, on the nature of these relations. Raz believes that they must expressly obtain between laws, whereas Bulygin believes that the connections between norms could be logically deducible; that is, they could be implicit: Carlos Alchourrón and Eugenio Bulygin, *Normative Systems* (Springer-Verlag 1971) 49. This distinction is not particularly important for the purposes of this thesis, so I do not consider it in further detail. For more on the deductibility/genetic distinction, see: Pablo Navarro and Jorge Rodríguez, *Deontic Logic and Legal Systems* (Cambridge University Press 2014) 144–49.

[20] HLA Hart, *The Concept of Law* (2nd edn, OUP 1994).

entirely new and unconnected legal system.[21] These secondary rules that produce the genetic structure are thus the bond shared by laws of a legal system. They ensure that the collection of laws referred to is produced by the same authority using the same rules. This is a significant enough commonality to merit designation as a legal system.

9.3.2. Relating Systems to Other Systems

The genetic and operative structures of a legal system tell us something about how laws relate to one another *within* that legal system, but it does not yet tell us anything about how legal systems relate to one another. This relation will necessarily be temporal; it is not usually the case that legal systems relate to one another for spatial reasons. At least, legal systems would usually not belong to the same *sequence* for spatial reasons. Along the temporal dimension, Raz distinguishes between 'momentary' and 'non-momentary' legal systems.[22] Momentary legal systems are those composed of norms that are valid according to the criteria for validity of a system as they existed at a particular moment.[23] Non-momentary systems, then, are chains of momentary systems related to one another by particular genetic relations.

Genetic structure is essential to understanding non-momentary systems[24] as it is this structure between momentary systems that accounts for the valid creation, change, and negation of laws and thus the organisation of the non-momentary system. Operative structure, by contrast, is particularly relevant to understanding momentary systems as it is through the lens of the operative structure that the punitive and regulative functions—the 'effect' of the laws existing at any given moment—come to light.[25] To put it in simple terms for present purposes: genetic structure

[21] John Finnis, 'Revolutions and Continuity of Law' in AWB Simpson (ed), *Oxford Essays in Jurisprudence: Second Series* (OUP 1972).

[22] Raz, *The Concept of a Legal System* (n 2) 187–97.

[23] ibid 34.

[24] ibid 184–85.

[25] ibid 185.

roughly maps to normative hierarchy in the non-momentary legal system, and operative structure reflects the actual normative claims of a particular momentary legal system.

9.4. LEGAL SYSTEM AND LEGAL ORDER

Raz's work established some important fundamentals of legal system theory. First, it established that legal systems are best thought of as relationships between sets. Second, he established that the genetic and operative structure of legal systems is essential in marking out the sets of laws that we actually count as legal systems. This is not yet enough on its own, however. Bulygin has further expanded on these fundamentals by introducing the distinction between legal systems and legal orders.

Legal systems, on this understanding, refer to a set of norms that is indexed to a particular point in time, and is thus analogous to what was above termed a 'momentary' system.[26] A legal order, by contrast, is a temporally ordered set of legal systems; a sequence of sets. Again, this is similar to what was above termed a 'non-momentary' legal system; however, it does build some further identity conditions into that concept that will be addressed below. For now, it suffices to note that under this distinction legal systems will have a short lifespan; they will change every time a norm is introduced or derogated in the system. Legal orders are what provides continuity between legal systems, and they can have much longer durations of existence. Indeed, they are the true objects of reference when legal officials refer to entities such as the 'South African legal system' or 'Canadian legal system'.

The first question we need to ask is one of identification: just what *are* legal orders? They are not just arbitrary sets of sets, so they cannot be identified solely by reference to their members like the ordinary mathematical sets above could.[27]

---

[26] Eugenio Bulygin, 'The Problem of Legal Validity in Kelsen's Pure Theory of Law' in Carlos Bernal and others (eds), *Essays in Legal Philosophy* (OUP 2015) 314.

[27] This is similar to the problem concerning the 'legal system' composed entirely of section 28 of all statutes. Such a set would be possible and would be defined extensionally by a list of its members, but it is not a plausible candidate for being considered a legal system or legal order.

So, what additional features distinguish them? Bulygin proposes the following five criteria:

> (1) The set of norms ($\mathcal{N}_1, \mathcal{N}_2, \ldots \mathcal{N}_n$) is the originating system (first constitution) of $\mathcal{O}_1$.
>
> (2) If a norm $\mathcal{N}_j$ is valid in a system $\mathcal{S}_t$ that belongs to $\mathcal{O}_1$, and $\mathcal{N}_j$ empowers an authority *x* to enact norm $\mathcal{N}_k$, and *x* enacts norm $\mathcal{N}_k$ at time *t*, then $\mathcal{N}_j$ is valid in the system $\mathcal{S}_{t+1}$ that belongs to $\mathcal{O}_1$.
>
> (3) If a norm $\mathcal{N}_j$ is valid in a system $\mathcal{S}_t$ that belongs to $\mathcal{O}_1$, and $\mathcal{N}_j$ empowers an authority *x* to derogate norm $\mathcal{N}_k$, which is valid in $\mathcal{S}_t$, and *x* derogates $\mathcal{N}_k$ at time *t*, then $\mathcal{N}_k$ is not valid in the system $\mathcal{S}_{t+1}$ that belongs to $\mathcal{O}_1$.
>
> (4) All valid norms in a system $\mathcal{S}_t$ that belongs to $\mathcal{O}_1$ that have not been derogated at time *t* are valid in the system $\mathcal{S}_{t+1}$ of $\mathcal{O}_1$.
>
> (5) All the logical consequences of valid norms in a system $\mathcal{S}_t$ that belongs to $\mathcal{O}_1$ are also valid in $\mathcal{S}_t$.[28]

As with much of Bulygin's writing, what this presentation gains in precision it loses through complexity. It will thus be aided by some additional narrative exposition. (1) extensionally defines some arbitrary set of norms that are the first constitution of the legal order. Essentially, these are the norms that a system just happens to begin with. (2) and (3) provide for the promulgation and derogation rules, respectively. The promulgation rules allow a properly empowered official to introduce new norms to the system, which creates a new momentary legal system. The derogation rules provide for the same, but through the deletion or subtraction of legal norms from the system. (4) provides that the validity of any non-derogated norms persists across legal systems within the same legal order. In other words, if a norm was valid in the system that occurred immediately prior to some act of promulgation or derogation, it will be valid in the momentary legal system that immediately follows (provided it was not the object of an act of derogation that caused the systems to shift). Finally, (5) provides logical closure to the legal order, providing that all the logical consequences of the norms in a legal system are valid in that system. This is a more idiosyncratic feature of Bulygin's own views on the contents of legal systems, and it is not particularly significant for present purposes.

---

[28] Eugenio Bulygin, 'Algunas Consideraciones Sobre los Sistemas Jurídicos' (1991) 9 Doxa 257, 263–64. This translation and representation is taken from: Navarro and Rodríguez (n 19) 200–01.

This understanding of legal systems and legal orders can be used to shed light on some of the more conceptual aspects of unconstitutionality. In particular, it is a helpful aid to considering validity in a more rigorous way. Navarro and Rodríguez have made some refinements to Bulygin's criteria for identifying legal orders above.[29] In particular, these are addressed to concerns concerning the promulgation of invalid norms. It seems clear that the identity for legal orders is determined in part by the legality of changes between legal systems (this is what (2) and (3) in the above quotation from Bulygin are directed to), but then what happens if a transition from one legal system to another obtains illegally (eg, through an unconstitutional addition or subtraction of law)?

One possibility is that illegal changes collapse legal orders, and so in each instance of an unconstitutional norm arising, the legal order changes. This proposal seems extreme, however. Implying constitutional revolution each time an unconstitutional norm is promulgated is not particularly parsimonious, to say the least, and it could potentially lead to difficulty in accounting for the continuity of legal orders.[30] Von Wright suggests, alternatively, that invalid norms do not imply changes in non-momentary legal systems at all, and are rather a recession from the normative system.[31] This has two potential problems, one particular to Bulygin's theory and one general: *(i)* it is not clear that Bulygin's theory could allow for the idea that unconstitutional norms are invalid (in the sense of being inapplicable) implied by von Wright (although they would certainly be 'invalid' in the sense that they are not members of a legal system within the legal order, it is not clear if there is any *normative* consequence that necessarily follows from this); *(ii)* it seems unrealistic to suggest that invalid norms do not change the legal system in some way if they are widely acted on and applied by legal officials.

---

[29] Navarro and Rodríguez (n 19) 201–05.

[30] Finnis, 'Revolutions and Continuity of Law' (n 21).

[31] '[T]he norm which is the result of the invalid normative act, by definition, does not belong to the system, but marks a recession from the system'. Georg Henrik von Wright, *Norm and Action: A Logical Enquiry* (Routledge & Kegan Paul 1963) 203

Problem *(i)* is soluble by noting that, for von Wright, the invalidity of a norm refers to the 'illegality of the act of issuing such a norm'.[32] The standard of validity in this case is the existence of another norm, not the *validity* of another norm. Given that this is a relative standard, a norm might be valid by reference to one higher norm, and invalid by reference to another. This sense of validity seems to refer to system membership, rather than being in force or applicable, and so does not provide a contradiction with Bulygin's theory.

More concerning, however, is problem *(ii)*. There is some intuitive appeal to the suggestion that insofar as an invalid norm can form a potentially widespread basis for legal action, it is part of the legal system somehow. However, this problem is in part animated by the idea that only valid and/or existing norms can be applicable to a legal dispute. Although this assumption seems even more foundational still, it is open to dispute. In the discussion below, I will attempt to motivate a greater focus on applicability, which allows for the idea that non-existent norms (in the sense of not existing *qua* law) could constitute the applicable norms to a legal dispute. If this is true, then both problems I posed to von Wright's analysis of invalid norms are solved, and the proposal that invalid norms effectuate no change to a legal system remains sound. However, this has to be interpreted alongside the view that invalid norms may still be *applicable* even if they are not themselves a part of the set of valid norms of the legal system. That is to say, even if invalid legal norms do not change the system by becoming incorporated into that system themselves (they are not members of the system) they may still be referred to and deemed applicable to particular cases that fall within the jurisdiction of that system.

9.4.1. Membership vs Inclusion

Although it seems reasonably clear what relationship momentary legal systems bear to prior or subsequent momentary systems, a point that is not so clear is the relationship that momentary legal systems bear to non-momentary systems. Drawing

[32] von Wright (n 31) 197.

on set theory, there are at least two possible relations that could obtain: *membership* or *inclusion*.[33] Membership implies that all momentary systems are themselves *elements* of an ordered set, which is the non-momentary system. Inclusion is a different relation; it implies that *both* non-momentary legal systems and momentary legal systems are sets, such that any given momentary legal system is a *subset* of some non-momentary legal system. This entails that both the momentary and non-momentary system will contain the same laws. Membership can be characterised as a relationship that obtains between containers and the things they contain, whereas inclusion is a relationship that obtains between containers (they may contain the same, or similar, things).

A brief example may help in understanding this concept. Take the number 6 and the set $\{6\}$. Now, 6 is a *member* of $\{6\}$,[34] and $\{6\}$ is *included* in $\{6\}$.[35] However, 6 is not included in $\{6\}$,[36] at least not in the sense of inclusion described above. This sense of inclusion would require that 6 be a subset of $\{6\}$, which it is not (as 6 is not a set or container). Figures 9.1 and 9.2 represent this difference in Venn diagrams.

---

[33] Bulygin, 'Time and Validity' (n 9) 173 fn 7.

[34] Mathematically, this would be represented as $6 \in \{6\}$.

[35] Mathematically, this would be represented as $\{6\} \subseteq \{6\}$. Every set is a subset of itself.

[36] $6 \nsubseteq \{6\}$. This is, of course, because 6 is not itself a set.

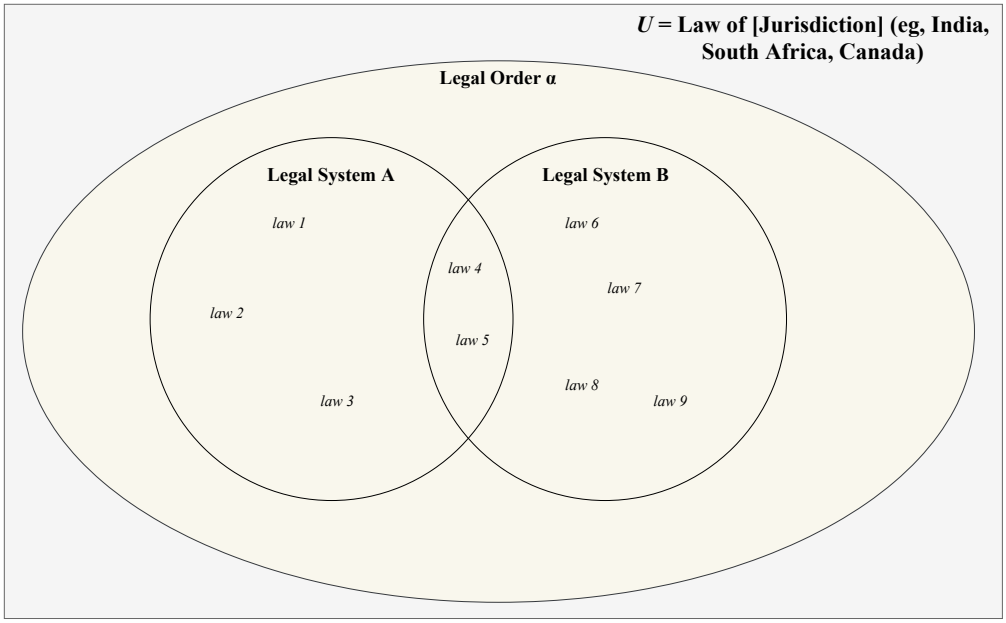


Figure 9.1: The Inclusion Model

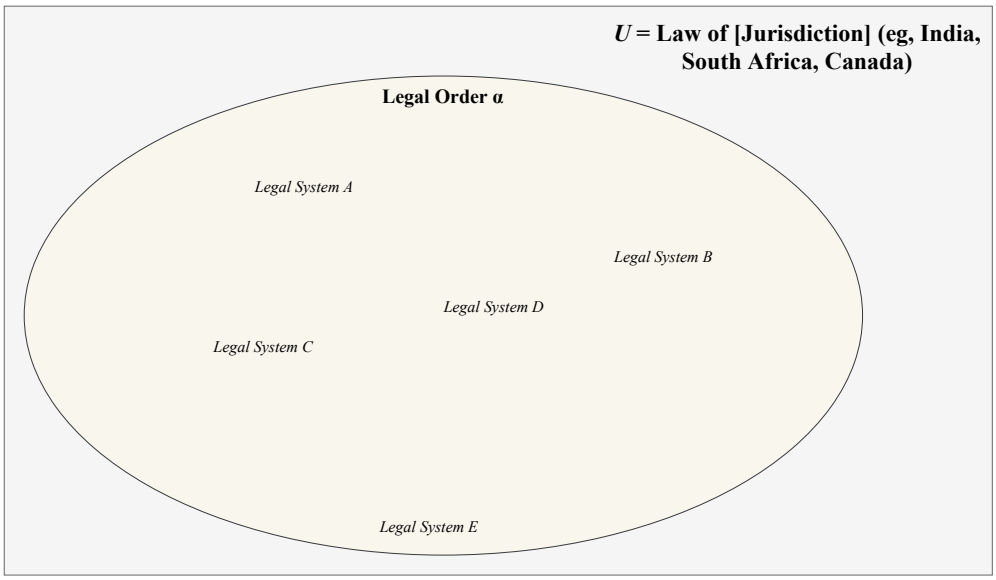


Figure 9.2: The Membership Model

There are two important things to consider here: *(i)* which relationship, membership or inclusion, better captures the relationship momentary legal systems bear to non-momentary ones? *(ii)* given that the membership reading will require that the non-momentary legal system be *ordered*, what does this mean and what further criteria must a set satisfy to meet this definition? The remainder of this section will consider the differences between the membership and inclusion relations, and then

it will consider the differences between ordered and unordered sets.

What, then, are the differences between inclusion and membership for the purposes of an analysis of legal systems? Each of these models has been further expounded upon by Navarro and Rodríguez, with a particular focus on a criticism of the inclusion model.[37] First, it is useful to observe a formal difference in the properties of membership and inclusion: inclusion is transitive, but membership is not.[38] Thus, where a momentary system is a *subset* of (ie, included in) a non-momentary system, then any members of the momentary system will be members of the non-momentary system. It is not true that where a momentary system is a *member* of a non-momentary system that any members of that momentary system are themselves members of the non-momentary system.

To see how this works, consider again a simple example with numbers. Take the numbers 2, 3 and 4, the set $\{2, 3\}$, and the set $\{2, 3, 4\}$. 2 is a member of $\{2, 3\}$. Additionally, $\{2, 3\}$ is included in $\{2, 3, 4\}$. We can therefore see that 2 is also a member of $\{2, 3, 4\}$. This is an application of transitivity. The same does not work with membership. Take, for example, the Finance Division (*fd*) in a company A. Say company A belongs to a global conglomerate Ω. Now, *fd* is a member of A, and A is a member of Ω, but it would be wrong to say that *fd* is a member of Ω; it is not the case that the Finance Division in company A is necessarily the same as the Finance Division in the conglomerate Ω.

Understanding this formal difference is crucial to understanding criticisms of the inclusion model. Navarro and Rodríguez raise two such criticisms. First, the inclusion view leads to the difficult conclusion that legal systems contain inconsistent sets of norms. Non-momentary systems will be mutually inconsistent, as they are achieved by the promulgation and derogation of new norms. Any set that contains

---

[37] Navarro and Rodríguez (n 19) 196–205.

[38] It is possible to demonstrate that membership is intransitive mathematically: consider the empty set $\varnothing$, as properties shared across members of the empty set are said to be *vacuously true*. Now consider a set $S : \{x, y, z\}$ where $x = \varnothing$, $y = \{\varnothing\}$ and $c = \{\{\varnothing\}\}$. Here it is true that $x \in y$ and $y \in z$ but $x \notin z$. So, the relation $\in$ (membership) on this set is not transitive. At the very least this proves that transitivity will be contingent.

all the contents of these non-momentary legal systems must therefore have contradictory elements. This is not as fatal as it may seem, though it is certainly a problem. That legal systems should be consistent—that is, free of internal contradiction—is a rationally desirable property of legal systems but not a necessary one.[39] Though it is not theoretically fatal, inconsistency is certainly undesirable, all other things being equal.

Second, and more importantly, the inclusion account provides a flawed understanding of what Navarro and Rodríguez call 'legal dynamics'. This is the problem of how change in legal systems is possible while simultaneously retaining the identity of that legal system. This is what I called the problem of *persistence* in the introductory section of this chapter. Consider the earlier point that a set's identity is a function of the identity of its members: if a non-momentary legal system is itself a superset of momentary legal systems, then any time a new momentary legal system is added to the set (because of the promulgation or derogation of a norm, for example) then the identity of the non-momentary legal system also changes. Because sets are static and not dynamic (they cannot change over time) the inclusion model fails to account for the dynamic element of non-momentary legal systems as it seems to entail that every time a law is promulgated or derogated we have an entirely new non-momentary legal system, as well as a new momentary one.

By contrast, the membership view has no issues here. Due to the conceptualisation of the non-momentary legal system as an ordered set, or sequence of sets, there is something other than the constituent members of the set that gives it its identity. The ordered set will also have some function that maps out the sequence of elements that are its members. It is this function (which will be whatever criteria mark out the static, momentary systems that are part of the non-momentary legal system) that gives the sequence its identity. Thus, it can be properly dynamic and retain its identity across time. The relevance of this to the temporal problems that come about with respect to unconstitutionality is that the membership view gives a much more

[39] See: Alchourrón and Bulygin (n 19) 62–63. The proof of this is quite mathematical and technical, and so it is omitted here.

plausible view of the persistence of laws. Its answer is that laws themselves do not persist across times, but rather *legal systems*—specifically, non-momentary legal systems—do. They can do this precisely because their identity is more than simply an array of norms and normative propositions, but rather it is a particular ordering of *sets of* norms and normative propositions achieved by a certain function.

Thus, it seems clear that the membership view (the view that non-momentary legal systems contain only momentary legal systems) is to be preferred over the inclusion view (the view that the non-momentary legal system contains both momentary legal systems and legal norms) for the reasons outlined above. It is to the membership view that the theory of legal systems advanced in this chapter subscribes. The discussion on ordered and unordered sets to follow further supports this stance.

9.4.2. Ordered and Unordered Sets

The term 'ordered' set was used to describe the non-momentary set under the membership reading. This naturally raises the question of what it means for a set to be ordered and how this differs from an unordered set. This is important for the temporal ordering of law within legal systems, as the relation that we will be considering to order the non-momentary legal system is a temporal relation. A set is *ordered* if a particular relation (for our purposes we shall use the two-place 'later than or at the same time as' relation) obtains between the members of the set and satisfies the following criteria:[40]

> *Reflexivity:* For all x in the set, x occurs later than or at the same time as x.
>
> *Antisymmetry:* For all x, y in the set, if x occurs later than or at the same time as y and y occurs later than or at the same time as x, then it entails that x is equal to y.
>
> *Transitivity:* For all x, y, z in the set, if x occurs later than or at the same time as y and y occurs later than or at the same time as z, then it entails that x occurs later than or at the same time as z.

[40] Eric Weisstein 'Totally Ordered Set', *MathWorld—A Wolfram Web Resource* ⟨http://mathworld.wolfram.com/TotallyOrderedSet.html⟩ accessed 12th October 2016.

These are the criteria for what are known as *partially ordered sets*. To make a set fully ordered, a further axiom must be added:

> *Comparability:* For all *x, y* in the set, either *x* occurs at the same time as or later than *y*, or *y* occurs at the same time as or later than *x*.

If the four criteria above are present, then a set is *totally ordered*. Returning to the idea of the non-momentary legal system as an ordered set of momentary legal systems, it is notable that this set of sets of legal systems satisfies the criteria for total ordering by the temporal relation 'later than or at the same time as':

> *Reflexivity:* Every momentary legal system within the non-momentary legal system occurs later than or at the same time as itself.
>
> *Antisymmetry:* For any two momentary legal systems in the non-momentary legal system, if they both occur later than or at the same time as each other, then they are equal.[41]
>
> *Transitivity:* If one momentary legal system in a non-momentary legal system occurs later than another, and this momentary system further occurs later than a third momentary system, then the first system occurs later than the third system.
>
> *Comparability:* One can take any two momentary legal systems in the non-momentary legal system and stand them in the relation that one occurs later than the other.

This shows that non-momentary legal systems can be ordered by the 'later than or at the same time as' relation.

The inclusion model (wherein the non-momentary legal system is a superset of all the momentary systems) cannot be ordered by the 'later than or at the same time as' relation, as the elements that would be ordered on this model would be normative statements or norms, rather than systems, and these would fail the antisymmetry criterion. It is not the case that if a normative statement (or norm) that occurs later than or at the same time as another normative statement is equal to that normative

[41] Of course in practice this can only be that the two systems occur at the same time, as it is contradictory to assert that *x* occurs later than *y* while simultaneously asserting that *y* occurs later than *x*. If the two systems occur at the same time, then they are identical, as the antisymmetry criterion requires.

statement. Thus, the inclusion model results in an unordered set (or, at least, a set that is not ordered by a two-place temporal relation).

To briefly take stock: we have seen that there is some attraction to characterising legal systems as sets. There are, however, two competing models of what type of set is at issue. The first of these is the *inclusion* model, which holds that the non-momentary legal system is simply a superset of all the momentary legal systems it contains, and which implies it contains all the elements of those sets (norms, normative statements, etc). By contrast, the *membership* model maintains that the non-momentary legal system is an ordered set of sets of momentary legal systems, which arranges those momentary systems by the 'later than or at the same time as' relation. The membership model was shown to be preferable to the inclusion model. Importantly, this means that the non-momentary system itself contains no norms or normative statements, it only contains an ordered sequence of legal systems.

9.4.3. The 'Internal and External' Time of Legal Norms

Another important concept highlighted by Bulygin in particular is that of the internal and external time of legal norms.[42] The interval between the time at which a norm is introduced and when it is derogated is the external time of that legal norm. For every interval of external time there corresponds at least one momentary legal system (this follows from how we have identified momentary legal systems as sets above). This ensures non-identity between successive legal systems, though in actuality successive legal systems will likely substantially overlap in content. The internal time of a norm, then, is the sequence of temporal moments during which that norm is applicable to some case.

The internal and external time of norms can be thought of as answering different questions: external time asks, 'during what time-interval(s) has the norm been included in the legal system?' and internal time asks, 'what time-interval(s) was the

[42] See generally: Bulygin, 'Time and Validity' (n 9).

norm intended to cover?'.[43] The external time of a norm is a function of system membership, and the internal time is a function of applicability.[44] Clearly, in order for this separation to fully make sense, there must be a distinction between system membership and applicability. This was already defended in chapter 8.

It is important to note the role temporality plays in law and legal systems here, and particularly the idea that the times during which a norm is a member of a legal system in the legal order (external time) may not necessarily be the same as the times during which that norm is applicable in the resolution of cases (internal time).

9.5. THE CONSISTENCY AND COMPLETENESS OF LEGAL SYSTEMS

This section briefly accounts for the consistency and completeness of legal systems, described in the introduction to this chapter. The concept of completeness of legal systems is of particular relevance because of how voidness *ab initio* seems to assume a set of complete and objective legal norms in a legal system. I reject an account of legal systems as complete in this way below, and I criticise accounts of law as being objective in chapter 10.

9.5.1. Raz's Error

I defined above what consistency and completeness mean for a legal system. Raz's model endorses the view that legal systems are complete, albeit with some restrictions. His earliest statement to this effect is quite unambiguous: '[a]ccording to every momentary legal system, every act-situation which is not prohibited by a specific law of the system is permitted.'[45]

More recently, and addressing the problem of legal gaps specifically, Raz has claimed that gaps necessarily arise where the law is indeterminate or yet to be defin-

[43] Rafaél Hernández Marín and Giovanni Sartor, 'Time and Norms: A Formalisataion in the Event-Calculus' (ICAIL '99 Proceedings of the 7th International Conference on Artificial intelligence and Law, Oslo, June 1999).

[44] Bulygin, 'Time and Validity' (n 9) 174.

[45] Raz, *The Concept of a Legal System* (n 2) 170.

itively resolved.[46] However, because of the above principle that everything that is not prohibited is permitted, no such gap can arise just because the law is silent.[47] Furthermore, Raz believes that this is required by an analytic truth about law: '[t]here are no gaps when the law is silent … closure rules, which are analytic truths rather than positive legal rules, [operate to] prevent the occurrence of [such] gaps'.[48] With respect to consistency, Raz acknowledges the possibility of legal conflicts[49] and so must be taken to reject the view that legal systems must be perfectly consistent.

So much for consistency, then. What of completeness? Raz seems to be of the view that legal systems *must* be complete. This is largely derived from his asserted equivalence between a lack of proscription and permission. I capture this by the phrase 'everything that is not prohibited is permitted'. This is the same equivalence that Dworkin uses in his argument to refute legal positivism.[50] Raz is similarly concerned to defend this view. It is therefore an important principle and worth critically analysing in more detail.

Bulygin has criticised Raz's equivalence between the lack of a proscription and a permission, on the basis that it can be reduced to a tautology and, thus, says nothing in support of the view that where law is silent it must also be complete.[51] This is because of an equivocation about what 'permitted' means above.[52] It could mean one of two things:

Strong Permission: There is a positive norm that permits ϕ.

---

[46] Joseph Raz, 'Legal Reasons, Sources, and Gaps' in Joseph Raz, *The Authority of Law: Essays on Law and Morality* (2nd edn, OUP 2009) 73.

[47] ibid 76–77.

[48] ibid 77.

[49] ibid 59–61.

[50] Ronald Dworkin, 'No Right Answer?' in Peter Hacker and Joseph Raz (eds), *Law, Morality and Society: Essays in Honour of HLA Hart* (Clarendon Press 1977).

[51] Bulygin, 'Kelsen on the Completeness and Consistency of Law' (n 12) 345–50.

[52] This criticism also builds on another distinction Bulygin draws, between norms and norm-propositions. Given that Raz accepts this distinction, I do not discuss it here. I return to this point below in the discussion on Bulygin's consistency and completeness theses.

Weak Permission: There is no norm prohibiting ϕ.[53]

Neither of these, however, can be substituted into the thesis that 'everything that is not prohibited is permitted' in a satisfactory way. If one adopts the strong permission view, then one gets 'everything which is not prohibited is positively permitted'. The truth of this will be contingent, and it certainly will not obtain in all cases. If one adopts the weak permission view, then one gets 'everything that is not prohibited is not prohibited' which, while tautologically true, is useless in any genuinely *explanatory* account of gaps in the law, and it is compatible with the existence of gaps in the law besides.[54]

This shows that there is not much hope to be had for the principle that 'everything that is not prohibited is permitted' as a closure rule guaranteeing the completeness of a legal system. I therefore suggest that Raz's argument for the completeness of legal systems should not be accepted.

9.5.2. Bulygin on Consistency and Completeness

Bulygin draws two distinctions to clarify concepts that he suggests have caused confusion on the completeness and consistency and legal systems. Recall that the most common argument for the completeness thesis in law is the thesis considered above to the effect that 'everything that is not prohibited is permitted'.

The first distinction Bulygin draws is between two types of deontic statement—norms and normative propositions:[55]

Norm: a command or prohibition.

Normative proposition: a descriptive proposition about a norm.

---

[53] Alchourrón and Bulygin (n 19) 121–27; Bulygin, 'Kelsen on the Completeness and Consistency of Law' (n 12) 345.

[54] Bulygin, 'Kelsen on the Completeness and Consistency of Law' (n 12) 346.

[55] Eugenio Bulygin, 'Norms, Normative Propositions, and Legal Statements' in Carlos Bernal and others (eds), *Essays in Legal Philosophy* (OUP 2015) 188–89.

The confusion between these types of deontic statements arises because they can be semantically identical but non-identical in terms of their descriptive-prescriptive dimension. Consider 'it is obligatory to remove one's hat indoors'. This can express either the speaker's endorsement of taking off hats indoors (in which case the sentence expresses a norm) or it can merely express the speaker *reporting* that there is a rule to that effect without endorsing it, in which case the statement expresses a normative proposition. This distinction is not only important because of the formal differences between norms and normative propositions (the former are not truth apt, the latter are), but because it is key to understanding how legal systems are not necessarily complete.

The second distinction to which Bulygin draws attention has already been deployed above in the discussion on Raz's arguments on legal gaps; the distinction between strong and weak permissibility in normative propositions:

> *Strong Permission:* There is a positive norm that permits ϕ.
>
> *Weak Permission:* There is no norm prohibiting ϕ.

I have already recapitulated Bulygin's argument for why neither of these understanding of permission in normative propositions can ground a view of law as complete, at least insofar as that view is founded on the view that 'everything that is not prohibited is permitted'.

The truth of the 'everything that is not prohibited is permitted' principle is open to only three interpretations, none of which compel the conclusion that a legal system cannot have gaps. On the norm interpretation, it is contingent (there may or may not be a positive norm issued by a legal authority that makes it true in a legal system). On the normative proposition readings it either results in an uninformative tautology (the weak reading) or a statement that must be false (the strong reading). Therefore, Bulygin rejects the view that legal systems are necessarily complete.

Only a brief note need be made on consistency: as with Kelsen[56] and Raz, Buly-

[56] A well-known feature of Kelsen's legal theory is his view that legal conflicts are impossible: 'No

gin is of the opinion that legal systems need not be consistent.[57] He builds this argument through criticising Dworkin and the early Kelsen's views that the law should be consistent, which were grounded in flawed ethical views. Particularly in Dworkin's case, the assumption that the moral principles injected into the law to guarantee its completeness and consistency will *themselves* be consistent is naive.[58]

9.6. CONCLUSION

Figure 9.3 summarises in a Venn diagram the model of legal systems endorsed in this chapter.[59]

---

conflict is possible between a higher norm and a lower norm … because the lower norm has the reason for its validity in the higher norm.' Hans Kelsen, *Pure Theory of Law* (Max Knight tr, University of California Press 1967) 208.
Kelsen was concerned that without at least some degree of hierarchical arrangement, legal systems would not exist *qua* systems. We would instead simply have sets of norms arranged in any order whatever, and therefore we would have no principled method to resolve conflicts of norms, as no one norm could claim priority over another. In this way, Kelsen viewed hierarchical arrangement as helping to secure *consistency* in a legal system.

[57] Bulygin, 'Kelsen on the Completeness and Consistency of Law' (n 12) 353.

[58] ibid 350–52.

[59] For presentational purposes only, this view is visually represented in a manner similar to the rejected inclusion view above. The summary diagram should therefore be viewed with the caveat that the endorsed view is the membership view and the relationship between the momentary and non-momentary legal systems in the diagram must be understood with this in mind.

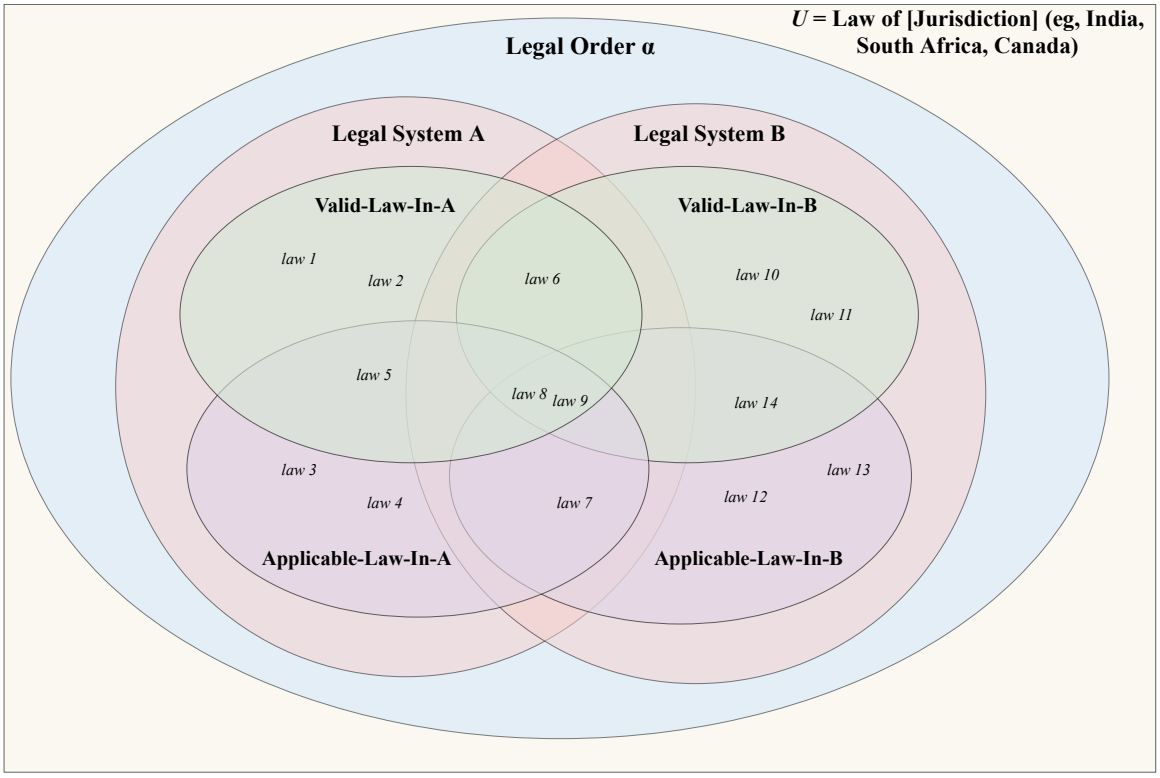


Figure 9.3: Model of a Legal System

It is important to note that applicable law is included within the sets marked as legal systems in figure 9.3. This may seem odd, as law that has been described as applicable but not valid includes foreign law in conflict of laws cases. One might think that such foreign law does not really 'belong' to the legal system. It must be recalled, however, that 'legal system', in the meaning of the diagram above, refers to a *temporary* or *momentary* legal system. This is a legal system that only exists for a short duration. Recall also that application is a complex relation that always requires a live case to be intelligible. Law can be 'valid' in the abstract, but it is only 'applicable' by reference to a concrete set of facts. In this way, the diagram above assumes to some degree that 'Legal System A' and 'Legal System B' are being considered relative to a particular set of facts. The main point is to accentuate how validity and applicability can overlap or come apart, and how they may both be relevant in considering how a 'legal system' might justify institutional action or juridical consequences in a particular set of facts.

It seems reasonably clear that Raz's original study left too much ambiguity on

important points such as the inclusion/membership problem discussed above. Although his model offers important insight on several points, the model latterly developed by Bulygin is more sophisticated in several respects. Most importantly, it is more developed in its account of the identity of non-momentary legal systems and how this identity relates to invalid legal norms. In general, the model advanced by Bulygin is preferable.

Regarding unconstitutionality, the analysis in this chapter supports the view that the temporal problems that can arise due to such declarations are best viewed as problems centred around the absence of a legal norm from a particular set of norms that are legally important. On an invalidity analysis, the problem centres around the absence of the norm from the set of member norms of a momentary legal system at a particular time or duration. On an applicability analysis, the problem instead centres around the presence or absence of the norm from the set of applicable norms at a particular time or duration.

Our understanding of legal systems also clarifies how it is that legal entities persist from one time to another; in other words, how legal entities can undergo change but maintain their identity across time. It is the legal system and not individual legal norms that persists. Legal norms have both their scope and content *essentially*, so if one altered a legal norm by making it proscriptive instead of permissive, for example, then one would not have changed that legal norm; one would have created an entirely new legal norm. The same logic would apply if one expanded or contracted the classes of cases to which the legal norm applies. The thing that gives law a sense of continuity—in the sense that what came before influences what is the case now—is the persistence of legal systems. Legal norms can, of course, exist over durations. They could be members of multiple momentary legal systems. Importantly, however, they cannot change and still be the same norm.

To summarise: this chapter endorses a set-theoretic model of legal systems. It supports the membership view over the inclusion view with respect to the question of how non-momentary legal systems relate to momentary ones. It is also a view that accepts legal gaps and underdetermination, rejecting the view that legal sys-

tems are either complete or consistent. Finally, individual momentary systems will themselves arrange legal norms in various subsets. The subset of valid norms is only one of these subsets. Relative to a particular case, the set of applicable norms in a momentary legal system may differ from this set of valid norms (for example, if it requires the application of foreign law). These are the general features of the legal system this thesis endorses.

10 | Flaws in Existing Theories of Unconstitutionality

CHAPTER OVERVIEW

10.1. Introduction 284

10.2. The Character of Law Problem 285

10.2.1. Metaphysical Objectivity in Law 286

10.2.2. Dworkin and 'Modest' Objectivity 287

10.2.2.1. The Impossibility of the Herculean Standard . . . 288

10.2.2.2. Dworkin on the Possibility of Judicial Disagreement 291

10.2.3. Moore and 'Strong' Objectivity 293

10.2.3.1. K-P Semantics 293

10.2.3.2. Criticism of K-P Semantics 295

10.2.4. Conclusions on Legal Objectivity 298

10.3. The Retrocausality Problem 299

10.3.1. Motivating Judicial Law-Making 300

10.3.2. Judicial Law-Making and Legislative Law-Making 303

10.3.3. Law-Making Power Implies Law-Derogating Power 305

10.3.4. The Tension between Invalidity and Judicial Law-Making . 306

10.3.5. Two Solutions for Retrocausality 307

10.3.5.1. Resolution One: Eternal Law 307

10.3.5.2. Resolution Two: 'Quantum Law' 309

10.3.6. Conclusion 312

10.4. Institutional/Moral Problems 313

10.4.1. General Justice at the Cost of Individual Injustice 314

10.4.2. Response of Other State Institutions to Suspension 316

10.4.3. The Rule of Law 318

10.5. Conclusion 320

10.1. INTRODUCTION

In one sense, the models of unconstitutionality that were constructed in chapter 8 are descriptive rather than normative. They are a reinterpretation of the practices of the jurisdictions studied in Part II, and thus they are generated using descriptive data. This chapter moves beyond this paradigm and considers grounds of criticism that might be deployed against the models. In other words, this chapter will consider the models as normative proposals, rather than reconstructions of practice. The criticisms considered here can be grouped under three broad headings: *(i)* the character of law problem; *(ii)* the retrocausality problem, and *(iii)* institutional/moral problems.

The first two problems relate to the void *ab initio* doctrine. This doctrine requires that the unconstitutionality pre-existed the adjudication, and that the judge was 'discovering' the true state of the law. Any view endorsing this pre-existence thesis requires that law is *metaphysically objective*. This is a difficult property to convincingly maintain, as I aim to argue below by reference to Dworkin and Moore's defences of metaphysical objectivity in law. These views are already difficult to accommodate in a theory of legal systems that rejects completeness, such as that endorsed in chapter 9. This view on completeness alone may be sufficient to reject these objectivity theories. This chapter bolsters the case for rejection of these theories by criticising them on their own terms.

The retrocausality problem shows how the void *ab initio* theory is inconsistent with a non-objective view of law. It is difficult to square any plausible account of judicial creativity with unconstitutional law being treated as 'void'. The void *ab initio* doctrine requires an answer to the 'origin' question that situates the cause of the unconstitutionality at a point in the past. In a system where the set of constitutional norms can change over time through judicial interpretation, invalidity from the point of origin of the unconstitutional legislation can become difficult to jus-

tify. A piece of legislation could be unconstitutional because of a contradiction with a constitutional norm that *itself* post-dates the legislation. This would make that legislation void from its supposed promulgation because of a contradiction with a norm that itself did not exist at the point of promulgation. This seems to suggest the facially absurd idea that future laws, which we do not yet know about (and which, arguably, do not yet even exist), can materially affect the constitutionality of present laws.

The institutional/moral problems refer to a set of concerns that are mostly levelled at particular answers or innovations in response to the harshness of voidness *ab initio*. These are ideas such as attributing prospective-only effect to unconstitutionality or suspending the effect of unconstitutionality. The benefits of these innovations are frequently cited in jurisdictions that have been faced with potentially severe and alarming consequences of a declaration of unconstitutionality.[1] Although such innovations have benefits, they come with some detrimental features also. These criticisms are not as severe as those that attach to the void *ab initio* theory, and the understanding of unconstitutionality argued for in this thesis can assist in designing further modifications to these theories that address these criticisms.

10.2. THE CHARACTER OF LAW PROBLEM

This section considers Dworkin and Moore's defences of metaphysical objectivity in law. As set out in the introduction, this is relevant for answers to the 'derivation' question from chapter 8 that maintain that the answers to legal questions pre-exist the determinations of those questions. This answer must defend some difficult views on the existence criteria of laws.

Taking Dworkin and Moore consecutively is useful, as they each endorse a different approach to metaphysical objectivity in law. Dworkin is a 'modest' objectivist, and Moore is a 'strong' objectivist. Difficulties in each theorist's proposals there-

[1] The Canadian *In Re Manitoba Language Rights* [1985] 1 SCR 721 (SCC) case and the Irish *A v Governor of Arbour Hill Prison* [2006] IESC 45, [2006] 4 IR 88 case are good examples of this.

fore show how these two alternatives should be rejected, and some more minimal objectivity (or indeed, subjectivity) in law is preferable, at least so far as unconstitutionality is concerned.

10.2.1. Metaphysical Objectivity in Law

To scrutinise the claims to objectivity in 'right answer' theories of law, we must first determine what 'objectivity' itself refers to. Leiter distinguishes between at least three different species of objectivity:[2] *(i)* metaphysical objectivity, which concerns questions of whether the existence of an entity or class of entities is mind-dependent with respect to a person or class of persons; *(ii)* epistemological objectivity, which concerns our ability to attain true knowledge of things that are metaphysically objective, and *(iii)* semantic objectivity, which concerns whether certain classes of propositions (scientific, aesthetic, legal, etc) are truth-evaluable (which requires that the referents of these propositions themselves be metaphysically objective). Right answer theses are fundamentally about how the truth or falsity of a legal matter relates to the cognitive processes and statements of certain officials involved in the legal community (most pertinently, judges). For this reason, they impute a claim of metaphysical objectivity.

Leiter then distinguishes several sub-categories of metaphysical objectivity that might be applied to law:[3]

1. Subjectivism: what seems right to the cognizer determines what is right.
2. Minimal objectivism: what seems right to the community of cognizers determines what is right.
3. Modest objectivism: what seems right to cognizers under appropriate or ideal conditions determines what is right.
4. Strong objectivism: what seems right to cognizers never determines what is right.

In the discussion below, I follow Leiter in reconstructing Dworkin as a modest

[2] Brian Leiter, 'Law and Objectivity' in Brian Leiter, *Naturalizing Jurisprudence: Essays on American Legal Realism and Naturalism in Legal Philosophy* (OUP 2007) 257.

[3] ibid 259.

objectivist, and Moore as a strong objectivist. Given that the options above proceed on a scale that proceeds from 'less objective' to 'more objective', it makes sense to take the best test case for metaphysical objectivity in law to be offered by advocates of the final two types of objectivity. Indeed, there are at least some good *prima facie* reasons to think that the law is objective in one of these two ways. Consider that in cases of unconstitutionality, the result can surprise everyone in the legal community. If the Supreme Court rules that a statute is unconstitutional, it is irrelevant if the rest of the judicial and legal community would have thought otherwise before the ruling; the statute is now unconstitutional. This seems to require at least minimal objectivity with a very small community of cognizers and, indeed, a community that will change over time as courts are reconstituted. Additionally, Coleman and Leiter observe that many forms of legal positivism that endorse a conventionalist thesis also endorse some version of minimal objectivity.[4] Only the final two types of objectivity, however, are sufficient to ground any kind of right answer thesis (as I have been using the term here). These forms of objectivity deny that truth can be contingent *only* on the mental states of an individual or community.

10.2.2. Dworkin and 'Modest' Objectivity

Insofar as Dworkin can be placed into a classification of metaphysical objectivity at all, it seems likely that he fulfils the criteria for Leiter's 'modest' objectivity. The difficulty with categorising Dworkin does not arise from any real ambiguity in what his views are; rather, it is his anti-metaphysical stance that presents a challenge. Dworkin maintained that debates over the facts that make a moral view 'true' or 'objective' are not really debates about what external fact, if any, makes that view true or false but are instead just straightforward ethical disagreement. This view has not been tremendously popular, it must be said. Leiter opines that Dworkin adopts it in the teeth of '[t]wo thousand years of metaphysics'[5] and Moore dubbed Dworkin's meta-

[4] Jules Coleman and Brian Leiter, 'Determinacy, Objectivity, and Authority' (1993) 142 U Pa L Rev 549, 612.

[5] Leiter, 'Law and Objectivity' (n 2) 272.

physical views the 'ostrich position'.[6]

So, putting Dworkin's own scepticism on metaphysics aside, it has been argued that he fits best within a paradigm of modest objectivity.[7] This is because the Herculean judge can be seen as an epistemologically idealised knower.[8] Whatever seems right and just to Hercules is the best theory of law. Most of the problems around this theory revolve around this idea of ideal epistemic conditions. We must ascertain what they are for a given domain and, even if we are able to do that, we might be forced to admit that in practice these conditions will never obtain. There are thus two epistemic hurdles: 'can we know what the ideal knower would look like?' and 'could we ever know what the ideal knower knows?'

10.2.2.1. The Impossibility of the Herculean Standard

Coleman and Leiter have supplied a tentative list of the kinds of qualities an epistemically ideal judge might have, in answer to the first of these questions:[9]

> The ideal judge must be:
>
> (1) fully informed both about
> (a) all relevant factual information; and
> (b) all authoritative legal sources (statutes, prior court decisions);
> (2) fully rational, eg, observant of the rule of logic;
> (3) free of personal bias for or against either party;
> (4) maximally empathetic and imaginative, where cases require, for example, the weighting of affected interests; and
> (5) conversant with and sensitive to information cultural and social knowledge of the sort essential to analogical reasoning, in which differences and distinctions must be marked as 'relevant' or 'irrelevant'.

It is striking that being a perfect moral reasoner is not on this list; after all, on some accounts legal reasoning is largely not autonomous and is just an applied form

---

[6] Michael Moore, 'Legal Reality: A Naturalist Approach to Legal Ontology' (2002) 21 Law & Philosophy 619, 632–35.

[7] Coleman and Leiter (n 4).

[8] See chapter 8 for a discussion of Hercules.

[9] Coleman and Leiter (n 4) 630.

of moral reasoning.[10] Judges will routinely have to grapple with moral issues implicated by the law, particularly in constitutional law when reference is made to abstract moral ideas such as 'equality' or 'justice'. It is not clear whether the criteria above could lead a judge to understand these concepts.[11] Coleman and Leiter's escape from this (and escape from the right answer thesis in general) is to maintain that their modest objectivity does not require total legal determinacy. That is, a judgment delivered under the above five conditions might still point to no legal fact.[12] It is not clear whether this rejection of determinacy allows for divergence of opinion among epistemically ideal knowers, but this seems unlikely.

Hercules is presented as having similar characteristics to the epistemically perfect imaginary judge above. He is possessed of 'superhuman skill, learning, patience and acumen'.[13] Whether Dworkin can fairly be classified as a modest objectivist depends on the role Hercules plays in Dworkin's argument in Hard Cases.[14] If Hercules' conclusions about law fix what the right answer to a legal dispute is, then Dworkin seems to fit this category.

Coleman and Leiter advance precisely this understanding of Dworkin's argument.[15] It is not entirely clear whether this was Dworkin's own intention. At one point, Dworkin suggests that the illustrative point of Hercules' method is that this perfect judge would be drawn to accept his right answer thesis.[16] If this is so, then Hercules could be taken as little more than a conjecture on how an ideal and perfect

---

[10] Perhaps the most well-known advocate of this sort of view is Raz: Joseph Raz, *Ethics in the Public Domain* (rev edn, OUP 1995) 326–40; Joseph Raz, 'Postema on Law's Autonomy and Public Practical Reasons: A Critical Comment' in Joseph Raz, *Between Authority and Interpretation* (OUP 2009) 376–81.

[11] The maximal abilities in empathy and imagination might suggest this but being able to empathise and imagine are different from being able to morally weigh options and decide on a course of action.

[12] Coleman and Leiter (n 4) 630, fn 157.

[13] Ronald Dworkin, *Taking Rights Seriously* (Duckworth 1977) 105.

[14] Ronald Dworkin, 'Hard Cases' (1975) 88 Harv L Rev 1057.

[15] Coleman and Leiter (n 4) 633–34.

[16] Dworkin, 'Hard Cases' (n 14) 1088, fn 23.

judge would prefer Dworkin's model of adjudication over others. Given how Dworkin's rights thesis works, however, and how it demands coherence in the law from the 'seamless web' metaphor,[17] it seems to be that it really does require Herculean judges, or at least judges that aspire to the Herculean ideal. If this is true then, as Coleman and Leiter contend, real life judges are required to try to decide cases as Hercules would.[18] This entails modest objectivity.

The chief problem with modest objectivity might be characterised as an access problem, or a problem of mapping epistemic ideals to an epistemically flawed reality. Assume that the 'real' law is accessible only to those possessed of Hercules' ability. If no judge in fact has this ability, does this not entail that no real judge can ever really know the true legal position? Coleman and Leiter suggest a resolution to this problem that distinguishes between *de jure* and *de facto* inaccessibility.[19] Something is *de jure* in accessible if on the terms of the theory of that thing, the thing is unknowable by humans. Something is *de facto* inaccessible if, as a matter of contingency, we do not know what the fact is, but we *could* come to know what it is. The issue for modest objectivity in law then is whether the position of the idealised legal knower is *de jure* inaccessible. In principle, at least, it is not.

In his own work, Leiter has pushed this problem further.[20] He has suggested that law cannot be property normative if it cannot be known, which seems to be implied by the *de jure* inaccessibility of law on both the strong and modest objectivity views:

> Part of the concept of law is that it is *normative* or *reason-giving*. Law cannot be normative, however, if *unknown*. This is why we need an answer to the question of epistemic access, for undetectable legal facts cannot *give* reasons, i.e., cannot be normative. Any conception of the law as Strongly or Modestly Objective raises the specter of the law being unable to fulfill its normative function, insofar as the specter of *de facto* inaccessibility seems a live one. Only a conception of the law as Minimally Objective is, it seems, guaranteed to be compatible with the normativity of law, precisely

[17] Dworkin, 'Hard Cases' (n 14) 1093–96.

[18] Coleman and Leiter (n 4) 634.

[19] ibid 631.

[20] In his work co-authored with Coleman they diverged on which type of objectivity they thought was preferable, with Coleman preferring modest objectivity and Leiter preferring minimal objectivity: ibid 627.

> because (1) communal consensus is constitutive of legal facts, and (2) such consensus is necessarily accessible to that community.[21]

Leiter's argument here is persuasive. The intuition that only things that we know about can form guides to action seems sensible, as does the view that communal consensus plays a constitutive role in legal facts. If this is true, then both modest and strong objectivity are seriously flawed.

10.2.2.2. Dworkin on the Possibility of Judicial Disagreement

A different strand of criticism considers whether Dworkin's right answer thesis is successful on its own terms, even if one were to grant all its premises. Kress, for example, has advanced what he calls the 'ripple effect' criticism of Dworkin's 'right answer' view.[22] I summarise this argument as follows:

> (1) In a hard case, a judge must adjudicate on a legal question (Q)
> (2) Q is not entailed by the legal principles of the system (P); it is *unsettled law*.
> (3) The judge must therefore decide whether Q or not-Q forms part of the soundest theory of the (settled) law (STL).[23]
> (4) Either Q or not-Q therefore becomes part of the settled law.
> (5) However, because Q or not-Q is now part of the settled law, the contents of STL are now also different.
> (6) If the contents of STL can differ in this way, then the order in which legal decisions are resolved by the courts matters, as STL evolves to react to new data that is generated by its own application to unsettled law.

Kress' objection relies on some important formal characterisations of Dworkin's right answer thesis. It requires the operation that relates the past law to the present law to be non-associative,[24] non-commutative,[25] or both; that is to say, Kress re-

---

[21] Leiter, 'Law and Objectivity' (n 2) 271. Emphasis original.

[22] Kenneth Kress, 'Legal Reasoning and Coherence Theories: Dworkin's Rights Thesis, Retroactivity, and the Linear Order of Decisions' (1984) 72 CLR 369, 380–81.

[23] This principle of bivalence, or the 'law of the excluded middle' endorsed by Dworkin's rights thesis here has been the subject of independent criticism: Joshua Geller, 'Truth, Objectivity, and Dworkin's Right Answer Thesis' [1999] UCL Jurisprudence Rev 83.

[24] An operation $*$ satisfies the associative law if $(a * b) * c = a * (b * c)$. If an operation does not satisfy the associative law, then the order in which the operation is carried out on the operands matters to the result of the operation.

[25] An operation $*$ satisfies the commutative law if $a * b = b * a$. If an operation does not satisfy the

quires that the evolution of law is path dependent. It is true that Dworkin himself has oscillated on this point. I discussed in chapter 8 some seemingly inconsistent statements on whether Hercules has subjective judgment and, if he does, whether it figures into his adjudication meaningfully.

It is at least a plausible option for Dworkin's theory of constructive interpretation to maintain that Herculean judging always pushes the law in a particular direction. This would be something like a perfect image or ideal of that legal system—the ideal that the Herculean judge is meant to identify and pursue. If the law is always judicially developed by this process, and only this process, then as far as judicial revision of the law, and retrospective application thereof, is concerned it is at least arguable that the order of input of 'data' into the judicial interpretation 'function' does not matter. If the function that is applied recursively (STL) requires the judge to use a set of principles that is coherent and consistent,[26] then it will ultimately generate the same legal system. Kress's objection implicitly requires that either a non-fixed set of principles is being applied, or that different sets of principles are being applied by different judges.[27]

It is significant that Dworkin himself later acknowledged the likelihood of judges coming to different conclusions in applying his interpretive method.[28] One might think that this confirms Kress's point, but Dworkin's concession is that although *in principle* there should be no divergence there often is because some judges will have mistaken views as to the best moral-political theory of the legal system. Again,

---

commutative law, then the order of the operands matters to the result of the operation.

[26] See also the discussion in chapter 9 on this issue; it is in fact highly unlikely that legal systems are completely consistent.

[27] It is worth noting briefly that this defence of Dworkin would also face some significant hurdles. The greatest of these would be that it implicitly holds that moral systems are totally complete and consistent systems of rules. This is unlikely. Leiter has noted that the prospects of objectively 'right' legal answers decline proportionately to the law's incorporation of moral criteria (Brian Leiter, 'Objectivity, Morality, and Adjudication' in Brian Leiter, *Naturalizing Jurisprudence: Essays on American Legal Realism and Naturalism in Legal Philosophy* (OUP 2007) 66-67).
Even aside from this point, philosophers of ethics have expressed doubts that morality works this way: Russ Schafer-Landau, 'Moral Rules' (1997) 107 Ethics 584.

[28] Ronald Dworkin, *A Matter of Principle* (Harvard University Press 1985) 17, 162.

assuming ideal Dworkinian conditions, it does not seem clear why order would matter in adjudications if every judge was a perfect Herculean reasoner. That they are not is no fault of Dworkin's theory, at least *qua* normative theory of interpretation. Dworkin can instead characterise this as weaknesses in the real world that fail to correspond to his Platonic ideal of judge and legal system.

10.2.3. Moore and 'Strong' Objectivity

Recall in Chapter 8 I drew attention to five theoretical points pointed out by Moore as characterising 'full-blooded' metaphysical realism (a view to which Moore himself subscribes). For a given class of entities, such a realist will maintain: (1) that the entities in question exist, (2) that they are mind-independent, (3) a correspondence theory of truth whereby the predicate 'is true' derives meaning from its correspondence to some independent state of affairs, (4) a truth-conditional theory of the meaning of sentences, whereby semantics is dependent on representation of how things actually are and (5) a causal theory of the meaning of words.[29]

10.2.3.1. K-P Semantics

For now I will focus on points (4) and (5), which relate to a particular semantics of law and, more importantly for the present context, semantics of judging that Moore endorses: the Kripke-Putnam theory of reference.[30] This theory holds that there are real meanings for words (that is, our language actually picks out certain features of the world) and it is *these* meanings of words that judges should seek out and apply in interpretive disputes.[31] More technically, it is a semantics that gives priority

[29] Michael Moore, 'The Interpretive Turn in Modern Theory: A Turn for the Worse?' (1989) 41 Stan L Rev 871, 878–79.

[30] I will refer to this alternatively as 'K-P semantics'.

[31] Michael Moore, 'A Natural Law Theory of Interpretation' (1985) 58 Cal L Rev 277; Moore's endorsement of K-P semantics emerges particularly clearly in his later work: Michael Moore, 'Law as a Functional Kind' in Robert George (ed), *Natural Law Theory: Contemporary Essays* (OUP 1992); Michael Moore, 'Justifying the Natural Law Theory of Constitutional Interpretation' (2001) 69 Fordham L Rev 2087.

to reference (the object the word picks out) over sense (whatever the term actually expresses) in determining the meaning of an expression.[32] An alternative 'conventionalist' semantics gives priority to sense over reference in determining meaning.

A typical example deployed in favour of K-P semantics is the meaning of the term 'water'. Rather than defining this according to the criteria for some concept 'water' (which might include criteria like 'transparent', 'liquid', etc) the direct or causal theory requires a stricter identity: water is H_2O. This is to hold that in *any world* where 'water' appears, its proper reference is the substance whose molecular arrangement is H_2O. This entails that in a world where, say, 'water' was the term used to refer to H_2O_2 (hydrogen peroxide), this would be an incorrect application of the term 'water'.

Whether one believes K-P semantics is intuitive on this point is not necessarily the most pressing point for the legal theorist. If there is *any* case for K-P semantics, it will be strongest with natural kind terms like 'water', which can be given by some essential feature that it and only it possesses. Other classic examples of terms with easily identifiable essences include 'gold' (Au)[33] and 'tiger' (which can be given by a certain configuration of DNA). The difficulty for the legal theorist is that law is unlike any of these. The most plausible candidate for an essential feature of law is the particular function(s) it performs. Indeed, few (if any) jurisprudential theorists think that positive law is a natural kind that cuts the world at the joints.[34] Given this, the legal theorist cannot simply be satisfied whether K-P semantics works with natural kind terms that might be considered its 'home turf'; the legal theorist must apply the theory to the more difficult case of law, a highly complex social artefact.

---

[32] Michael Moore, 'Justifying the Natural Law Theory of Constitutional Interpretation' (2001) 69 Fordham L Rev 2087, 2091.

[33] More precisely, it would be 'the element that has 79 protons'—nothing can possess this property and not be gold. Differences in electrons can occur (gold could be differently ionised) and differences in neutrons can also occur (gold has several isotopes).

[34] Brian Leiter, 'The Demarcation Problem in Jurisprudence: A New Case for Skepticism' (2011) 31 OJLS 663, 666.

10.2.3.2. Criticism of K-P Semantics

The legitimacy of extending K-P semantics to artefactual entities like law is a vexed question, precisely because of the doubt over whether such artefactual kinds have any uniquely identifying essential features.[35] More modern commentary has tended to disfavour the application of K-P semantics to artefactual kinds, such as law.[36] In particular, Leiter has emerged as a strident critic of this view.[37] He has three central avenues of criticism:

(1) K-P semantics might, itself, be philosophically unfounded or flawed;

(2) Law does not have an essence that would fix the meaning of the expression 'law' (this also applies to substantive legal expressions like 'statute' and moral expressions within the law, such as 'equality');

(3) Legal interpretation must be more than just a correct account of meaning, as it also requires non-semantic elements—for example, statutory interpretation is about giving effect to legislative intention that may or may not coincide with the true meaning of terms used in a statute.

The first point is not amenable to decisive resolution, and so I do not consider it further; it is, however, certainly a possibility. The second point is, to my mind, the strongest and if it is made out then the third does not arise for consideration. Debate about law's social artefactual nature has become more rich in recent years,

---

[35] Some doubt this: Stephen Schwartz, 'Putnam on Artifacts' (1978) 87 Philosophical Review 566; Stephen Munzer, 'Realistic Limits on Realist Interpretation' (1985) 58 S Cal L Rev 459; Dennis Patterson, 'Realist Semantics and Legal Theory' (1989) 2 Can J L & Jurisprudence 175. Others do not: Michael Moore, 'The Semantics of Judging' (1981) 54 S Cal L Rev 151; Dennis Patterson, 'Semantics and Legal Interpretation (Further Thoughts)' (1989) 2 Can J L & Jurisprudence 181.

[36] See, for example, Nicos Stavropoulos, *Objectivity in Law* (OUP 1996) and even Moore himself in later writing has preferred to style terms such as 'contract' and 'malice' as defined by the purpose to which they are put, rather than their nature or function: Michael Moore, 'Truth-makers for Propositions of Law and for Propositions About Law' in Geert Keil and Ralf Poscher (eds), *Vagueness and the Law* (OUP 2014).

[37] See generally: Leiter, 'Objectivity, Morality, and Adjudication' (n 27); Leiter, 'Law and Objectivity' (n 2).

but one near-constant feature in that literature is that if law as an artefact has anything essentially, it is certain *functions*.[38] The more difficult point is identifying what those functions are, and how we discriminate between essential and non-essential functions.

One key difference, as already intimated above, between natural and artefactual kinds is that natural kinds have their natural properties *essentially* whereas artefactual kinds do not have their natural properties essentially. Water and hydrogen peroxide are a good example here again: they have the natural properties of 'being H_2O' or 'being H_2O_2' essentially. There are many differences in their natural properties that can be observed that stem from this essential difference: heat both substances to 100 degrees Celsius and one will explode, and one will not.[39] However, artefacts have non-essential natural properties as well; a table made of metal is as much a table as one made of wood. As I said above, artefacts have proper functions that they are intended to perform. Money is more than just pieces of brightly coloured paper because it is intended to perform a certain function in providing value for transaction. Let us suppose that artefacts at least have their intentional functions essentially. Positive law would still present difficulty in this regard because it performs a gamut of regulative functions that are not easily reducible to one abstract, general formula for the 'function' of law. This has led some to doubt the possibility of law having an essence or nature at all.[40] Green has recently responded to this objection by focusing on the intentions of the creator of the artefact.[41] He gives the example of printer drivers, which are strings of code written to drive printers. If they do any-

[38] Neil MacCormick, *Institutions of Law: An Essay in Legal Theory* (OUP 2007); Jonathan Crowe, 'Law as an Artifact Kind' (2014) 40 Monash U L Rev 737; Luka Burazin, 'Can There Be an Artifact Theory of Law?' (2016) 29 Ratio Juris 385; Brian Leiter, 'Legal Positivism about the Artifact Law: A Retrospective Assessment' in Luka Burazin, Kenneth Himma and Corrado Roversi (eds), *Law as an Artifact* (OUP 2017).

[39] The answer here will be apparent to anyone who has ever used a kettle. Peroxide vapour can detonate at temperatures higher than 70 degrees Celsius.

[40] Leiter, 'The Demarcation Problem in Jurisprudence: A New Case for Skepticism' (n 34) 663.

[41] Leslie Green, 'The Morality in Law' in Luis Duarte d'Almeida, James Edwards and Andrea Dolcetti (eds), *Reading HLA Hart's* The Concept of Law (OUP 2013).

thing else, they are not printer drivers. But this example does not, I think, do the work Green intends for it. Say a piece of code drives a particular printer. But say that, through some quirk in the way definitions were set out in the code, it can also drive certain microphones. Is it any more just a printer driver?[42] Perhaps it is both entities now, or just one depending on the purpose for which it is actually used. The point is this: design-features are not so determinative as Green would have them be.

Additional difficulties arise because artefacts may never perform their function, or may even outright malfunction, and maintain their identity as the particular artefact in question.[43] If I have an old mobile phone that can no longer send texts or make calls, but which serves perfectly well as a doorstop, it does not simply cease being a mobile phone. Additionally, artefactual kinds may have accidental as well as intentional functions.[44] Again, deciding which of these functions to give priority to is a challenging task.

One final objection to strong objectivity of the Moorean stripe is the so-called 'global error' objection.[45] Effectively, strong objectivity seems to entail the view that the entire community can be genuinely mistaken as to the correct legal position. This seems very implausible given that 'what the law is' is part-determined by what a relevant class of officials (which could include judges, lawyers, politicians, depending on how wide one cast the net) think the law is. The possibility for *global* error seems to be ruled out on any account that accepts some role of the mental states of at least some people in the constitution of law. Most accounts accept some role of

---

[42] Brian Leiter, 'Legal Positivism about the Artifact Law: A Retrospective Assessment' in Luka Burazin, Kenneth Himma and Corrado Roversi (eds), *Law as an Artifact* (OUP 2017) A similar line of reply is advanced in:

[43] Lynne Rudder Baker, *The Metaphysics of Everyday Life: An Essay in Practical Realism* (Cambridge University Press 2007) 55–57.

[44] Wybo Houkes and Pieter Vermaas, 'Actions Versus Functions: A Plea for an Alternative Metaphysics of Artifacts' (2004) 87 The Monist 52.

[45] Andrei Marmor, *Interpretation and Legal Theory* (OUP 1992) 85–102; Matthew Kramer, 'Is Law's Conventionality Consistent with Law's Objectivity?' (2008) 14 Res Publica 241; Brian Bix, 'Global Error and Legal Truth' (2009) 29 OJLS 535; Brian Bix, 'Metaphysical Realism and Legal Reasoning: The Philosophy of Michael S Moore' in Kimberly Ferzan and Stephen Morse (eds), *Legal, Moral, and Metaphysical Truths* (OUP 2016).

this type.

10.2.4. Conclusions on Legal Objectivity

There are, thus, difficulties with both Dworkin and Moore's strategies in advancing a right answer thesis in law. It is worth noting that one of the uncomfortable aspects of both theories is that they prescribe *unique* right answers. It is much simpler to say that there is a *range* of 'correct' or 'acceptable' legal answers to a problem and choice between these is arbitrary, or contingent on other factors. This steers away from the equally undesirable extreme realist view that there are no a priori correct or incorrect legal answers on any matter until a judge has authoritatively ruled on it.[46]

Rosati has drawn attention to a difficulty that seems to be shared by both the modest and strong objective views.[47] She takes the example of the sentence 'The law is *Plessy*, but *Plessy* was wrongly decided'.[48] It is a puzzling feature of legal practice that we can maintain this sort of view without contradiction. If we assume that judges are bound to apply the law, then the law is not *Plessy* only if it is decided in a way that is contrary to the true legal facts. In other words, if *Plessy* was wrongly decided, then it is not the law.

One way to resolve this apparent contradiction, according to Rosati, is to say that the first part of the sentence 'the law is *Plessy*' merely summarises whatever it is that judicial authority has had to say on the matter, but the second clause '*Plessy* was wrongly decided' expresses a judgment as to what the law really is.[49] But suppose *Plessy* was never overturned and, indeed, will never be overturned. It might be fine to say that nobody in this universe ever discovered the true requirements of morality under those circumstances, but it seems decidedly more odd to say that people

[46] Anthony Woozley, 'No Right Answer' in Marshall Cohen (ed), *Ronald Dworkin and Contemporary Jurisprudence* (Duckworth 1984).

[47] Connie Rosati, 'Some Puzzles About the Objectivity of Law' (2004) 23 Law & Philosophy 273.

[48] Rosati (n 47) 291, citing *Plessy v Ferguson* 163 US 537 (1896).

[49] Rosati (n 47) 292.

never knew what the law was and will never know. This seems to be deeply tied to intuitions about law as an *artefactual* normative system.

Law's artefactual nature is perhaps the most general heading under which complaints against objectivity theories can be grouped. Right answer and objectivity theories are insufficiently attentive to the artefactual nature of law and attempt to situate it with natural kinds or highly abstract moral systems. Law is too intimately related to community practice to admit of complete explanation by reference to these ideas, and therefore strong and modest objectivity theories fail. A theory of unconstitutionality that commits itself to legal systems having a complete set of unique right answers to constitutional questions should be rejected.

10.3. THE RETROCAUSALITY PROBLEM

This section presents a criticism of the alternative prospect for the void *ab initio* theory of unconstitutionality. If, as argued above, objectivity in law is untenable, then the void *ab initio* theory must either be compatible with subjectivity, or it must be discarded. Subjectivity in this sense refers to there being no fixed and objective answers to legal questions; that is to say, law requires some measure of creativity from the judiciary in cases where the law is indeterminate or there is a legal gap.

I aim to show that the following two propositions must produce a contradiction: *(i)* judges change the law through the creation of legal norms; and *(ii)* invalid laws are non-existent and void *ab initio*. The first of these is a thesis that could be embraced on any of the models of unconstitutionality endorsed here, as it is not a proposition about the nature of unconstitutionality. It may seem facially incompatible with the 'legislative nullity' answer to the derivation question, but it must be remembered that the derivation question is actually quite narrow. It asks only what the ground of the *unconstitutionality* in a case is. A judge could have a role in shaping the norms that result in unconstitutionality without necessarily having a constitutive role in the unconstitutionality itself.

The second of these propositions (the invalidity thesis) is a descriptive propos-

ition for these purposes. It is clear from the analysis in Part II and chapter 8 that unconstitutional law is often rendered invalid and void. The discussion of judicial creativity has not been included as a descriptive element of the legal systems considered in this thesis and so there is first some work done in the following subsections to motivate acceptance of judicial law-making. Once this is complete, I show how judicial creativity (subjectivity) and voidness *ab initio* contradict one another. This contradiction can only be resolved through the acceptance of one of two extreme further commitments. I therefore suggest that voidness *ab initio* is incompatible with a subjective account of law's dynamics.

10.3.1. Motivating Judicial Law-Making

It is increasingly well-accepted, among both academics[50] and the judiciary,[51] that judges have a law-making function that is distinct from that of the legislature, but still operates to generate new norms in the legal system.[52] I argue below how it makes good sense to consider that judges make constitutional law.

In part, this view of law relates to the repudiation of legal objectivity earlier in

---

[50] Raz is an interesting example here. One might suppose that his Sources Thesis requires the view that judges should apply only those legal norms that are validated by sources. Raz, however, thinks that the Sources Thesis in fact requires that judges engage in an exercise that is at least part-creative when they interpret the law. Interpretation arises, *inter alia*, where there are legal gaps, and where these arise in the law the courts must apply some other norm that, by virtue of *stare decisis*, results in the creation of new legal norms: Joseph Raz, 'Legal Reasons, Sources, and Gaps' in Joseph Raz, *The Authority of Law: Essays on Law and Morality* (2nd edn, OUP 2009); Joseph Raz, 'Law and Value in Adjudication' in Joseph Raz, *The Authority of Law: Essays on Law and Morality* (2nd edn, OUP 2009). Gardner accepts a similar view: John Gardner, 'Legal Positivism: $5\frac{1}{2}$ Myths' in John Gardner, *Law as a Leap of Faith* (OUP 2012) 34–35. Indeed, Raz has gone further and suggested that common law courts have an implicit power to change the law where it is asynchronous with justice or morality: Joseph Raz, 'Postema on Law's Autonomy and Public Practical Reasons: A Critical Comment' in Joseph Raz, *Between Authority and Interpretation* (OUP 2009) 377.

[51] *A v Governor of Arbour Hill Prison* (n 1); *Canada (Attorney General) v Hislop* 2007 SCC 10, [2007] 1 SCR 429.

[52] This is so even among modern proponents of views close to Blackstone's, such as Justice Scalia:

> I am not so naive (nor do I think our forebears were) as to be unaware that judges in a real sense 'make' law. But they make it *as judges make it*, which is to say *as though* they were 'finding' it—discerning what the law *is*, rather than decreeing what it is today *changed to*, or what it will *tomorrow* be.

James B Beam Distilling v Georgia 501 US 529, 549 (1991).

this chapter. If there are not objective truths about law, then legal officials must make those truths. The most obvious candidate legal officials in this case are judges. Judges may be constrained to a greater or lesser degree in making these truths; I am not taking a view on the degree of discretion afforded by the judicial role. However, if law is not fully objective then there must be *some* measure of discretion for judges to meaningfully address legal gaps. This implies some measure of law-making power, but that power might be materially different to *legislative* law-making. This nuance helps to see how judicial law-making *simpliciter* might not breach the separation of powers; rather, it would be the judiciary making law *in the same way the legislature does* that would be of concern.

One type of law that it seems judges often create is *constitutional* law. There are at least two types of claim that might support this: one is a conceptual claim about the nature of interpretation, and the other is an empirical claim about the practice of constitutional law.

Let us first consider the conceptual point first. Raz argues that it is a general truth regarding the concept of interpretation, irrespective of context, that it has a Janus-faced aspect: it is both backward-looking in elucidating what the law *is* and forward-looking in considering how the law *ought* to be.[53] If this is true of interpretation in general, then it is true of legal interpretation in particular. Constitutional interpretation gains its double-sided quality from the backward-looking impetus to ground constitutional interpretation in sources of constitutional authority, and the forward-looking impetus to give weight to current and future moral considerations.[54] If interpretation requires something more than simply reading what is in the text and necessitates *adding* meaning to the text, then interpretations of the Constitution add content to constitutional law that was not there in the pre-interpretative state.

---

[53] Joseph Raz, 'On the Authority and Interpretation of Constitutions: Some Preliminaries' in Larry Alexander (ed), *Constitutionalism: Philosophical Foundations* (Cambridge University Press 1998) 177.

[54] ibid 178.

Let us now consider the empirical claim.[55] Put bluntly, the claim here is that it is an observable fact in legal systems that more constitutional law accrues over time. As Gardner put it in one context, there is more constitutional law now than there was in the past.[56] This constitutional law must come from *some* source, and not all of it comes from formal amendment. The answer here, then, points to judges. This in turn trades on a distinction between the constitutional *text* and constitutional *law*.[57] Constitutional law is the accretion of the rules of the *acquis constitutionnel* as it develops through case law and judicial pronouncement. Either the constitutional text or the constitutional law is a valid object of interpretation. Though Gardner does not pass comment on which (if either) is, or should be, the proper object of constitutional interpretation, it seems to make sense that as a legal system ages and sees more constitutional litigation it is more likely that interpretation of constitutional *law* will arise more frequently. It seems best, however, to conclude that both are valid objects. The key point for present purposes is that if interpretation requires innovation, and the Constitution requires judges to interpret it, then it seems that the conclusion that judges make constitutional law is inescapable. This seems supported by both the conceptual and empirical claims here.

It is worth noting here that these claims regarding constitutional interpretation and creativity also serve to supplement the conceptual arguments against the objectivity of law above. I argued in the first section of this chapter that it is theoretically difficult to argue that law *by its nature* is objective. However, even if this claim were to be false or rejected, it could still plausibly be claimed that constitutional law *specifically* is an unlikely type of law to be objective. Voidness *ab initio* seems

---

[55] It is true that a holder of a Blackstonian view of precedent would likely reject the empirical claim: it seems sensible to think that a Blackstonian would be a kind of originalist about constitutional interpretation. See, for example, the comments of Scalia J to the effect that the United States constitution 'does not change from year to year' (*American Trucking Associations v Smith* 496 US 167, 201 (1990)). This, however, denies any creative aspect to the practice of interpretation and is, as such, an impoverished theoretical account of interpretation and is thus vulnerable to the conceptual argument above.

[56] John Gardner, 'Can There be a Written Constitution?' in Leslie Green and Brian Leiter (eds), *Oxford Studies in Philosophy of Law: Volume 1* (OUP 2011) 191.

[57] ibid 192.

to require that all constitutional law be objective; this entails that if even *some* unconstitutional law is not objective, then that theory fails. Many cases that result in declarations of unconstitutionality are arguments over contested concepts and values. To take just a few examples from the jurisdictions studied in this thesis, subjects such as abortion rights,[58] the rights of refugees,[59] the rights of same-sex couples,[60] the right to die,[61] and the right to compensation following compulsory acquisition[62] have invited constitutional consideration.

Even if some law is more objective than other types of law, it seems plausible to think that the constitutional law leading to declarations of unconstitutionality is often quite subjective. Many constitutional values are relatively objective (the colours of state flags, for example). Some of these more objective values may lead to unconstitutionality (eg, a referendum that did not follow constitutional procedure). However, most cases of unconstitutionality involve human rights of the kind given in the examples above. These are likely not objective criteria (full consideration of the debate over moral realism and moral absolutism is beyond my remit in this thesis).

10.3.2. Judicial Law-Making and Legislative Law-Making

If judges are to make law legitimately, in a manner consistent with the separation of powers, then they must be engaged in a practice that is distinguishable from legislation, as it is common case that judges cannot legislate. Gardner proposes a difference between judicial law-making and legislative law-making on the basis that judges create new legal norms by a special process of *legal reasoning* or 'reasoning-

[58] *Planned Parenthood of Southeastern Pennsylvania v Casey* 505 US 833 (1992); *Whole Woman's Health v Hellerstedt* 579 US — (2016).

[59] *Laurentiu v Minister for Justice, Equality and Law Reform* [1999] 4 IR 26 (SC).

[60] *National Coalition for Gay and Lesbian Equality v Minister of Home Affairs* [1999] ZACC 17, (2000) 2 SA 1; *Reference Re Same-Sex Marriage* [2004] 3 SCR 698 (SCC).

[61] *Carter v Canada (Attorney General)* 2015 SCC 5, [2015] 1 SCR 331.

[62] *Mahendra Lal Jaini v State of Uttar Pradesh* [1962] INSC 303, [1963] Suppl SCR 912.

according-to-law'.[63] The distinction works on the following basis: legislative law-making is characterised by the ability to make new legal norms on *entirely* non-legal grounds whereas judicial law making must include at least one non-redundant legal premise in its derivation of new law.[64]

It is important to note that legal reasoning on this scheme is reasoning that 'has already-valid legal norms among its major or operative premises, but combines them non-redundantly in the same argument with moral or other merit-based premises'.[65] Thus, the general scheme of legal reasoning should follow something like the following argumentative structure:[66]

> *Major premise(s)*: Legal Norm L_1.
> *Major/minor premise(s)*: Moral Norm M_1.
> *Conclusion*: Legal Norm L_2, derived from a relation of L_1 to M_1.

This implies, of course, that legal reasoning is related to moral reasoning. In fact, the distinguishing mark between legal and moral reasoning is that legal reasoning simply involves at least one operative and non-redundant legal premise. There is nothing in the difference with regard to the method. This requires that legal reasoning is just a subspecies of moral reasoning; indeed, it is only fully autonomous as a species of reasoning where morality runs out.[67]

On the basis of the above, it might be suggested that rather than the vague line of substance or principle that courts often attempt to draw between legitimate or illegitimate policy changes, a better test would be one of form such as the one posited above. So, judges are confined to including at least one legal premise in their derivation of new law, whereas the legislature is not so constrained.

---

[63] Gardner, 'Legal Positivism: $5\frac{1}{2}$ Myths' (n 50).

[64] ibid 40–41.

[65] ibid 39.

[66] A very similar point is also made by Bulygin: Eugenio Bulygin, 'Judicial Decisions and the Creation of Law' in Carlos Bernal and others (eds), *Essays in Legal Philosophy* (OUP 2015).

[67] Joseph Raz, *Ethics in the Public Domain* (rev edn, OUP 1995) 340.

10.3.3. Law-Making Power Implies Law-Derogating Power

If judges can create law, they can also derogate it. The derogation of a legal norm seems to require at least legislating for the negation of that norm, and possibly the enactment of a new replacement norm. Bulygin is unequivocal that this is precisely what judges do:

> The much debated problem of whether judges 'create' law or only apply it can be settled in favour of the first thesis, at least in the sense that they modify the legal system by imposing an ordering on its elements when they have to resolve contradictions, *disregarding some of the norm-contents (which amounts to derogating them)*.[68]

This seems to render clear that judges have a creative role, and that it is strongly linked to derogation. Bulygin further explains the logic of derogation here:

> Derogation, which leads to a new system … compris[es] two components, the act of rejection and the operation of subtraction. The act of rejection identifies a derogandum [set of expressly rejected propositions], and the resulting system is the remainder after subtracting a derogans [subset of the legal system which minimally contains all the same propositions of the system but without those expressly rejected] (corresponding to the derogandum) from the original system.[69]

The important point to note here is the operation of subtraction. Subtraction is logically the equivalent of adding a negative. $4 + 2$ and $4 + (-2)$ are both equal to 2. This logical relation between addition and subtraction demonstrates how law creation and derogation powers are two sides of the same coin. Therefore, accepting that judges make law implies that judges have a power to derogate law. Unconstitutional law, because of its incompatibility with the most supreme laws of the legal system, is a natural candidate for such derogation through judicial review.[70]

---

[68] Eugenio Bulygin, 'The Expressive Conception of Norms' in Carlos Bernal and others (eds), *Essays in Legal Philosophy* (OUP 2015) 166. Emphasis added.

[69] ibid 162.

[70] Recall, however, how it was discussed earlier in chapter 8 that derogation does not necessarily imply repeal, and that judicial review is not best conceptualised as a kind of repeal power.

10.3.4. The Tension between Invalidity and Judicial Law-Making

The difficulty in asserting invalidity and voidness as a consequence of unconstitutionality, while simultaneously reserving a creative role for the judiciary in shaping constitutional norms, may now be made clear. The difficulty lies in using a changing constitutional standard to assess the validity of laws by reference to a rule of automatic retrospective invalidity. This is the retrocausality problem, represented by the diagram below (where T_1 and T_2 are two different instants in time, T_1 is earlier and T_2 is later; CT is the constitutional text; CN_x refers to constitutional legal norm x; and LN_y refers to legislative legal norm y):

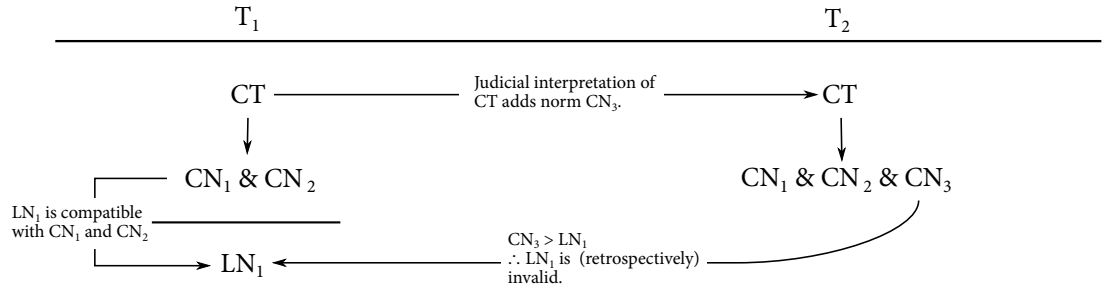


Figure 10.1: The Retrocausality Problem: Invalidity and Changing Law

As figure 10.1 shows, it is possible to derive the unusual conclusion that future laws in some sense 'cause' past laws to be invalid. It is important to stress that this is a matter of temporal perspective. From the perspective of an observer or participant in the legal system at T_2 above, the practice can be rationalised as the legal facts of the present having retrospective effect. From the point of view of an observer or participant in the legal system at T_1, however, things look different. For such a hypothetical person to accept the invalidity of the legislative norm at issue, they would have to be able to think that some (unknown to them) judicial determination in the future will add content to the constitution *and* that this content grounds the invalidity of the legislative norm.

Naturally, lawyers and laypeople alike tend to favour the present-tense perspective. This is all the more true when confronted with a pressing and immediate legal problem. We cannot change the past, but we can shape the present and, in turn,

hope to nudge the future along a particular path. This does not take away from the problem that the hypothetical perspective of a past person is an issue in the above theoretical model. Even a person thinking in a present-tense perspective needs to be able to accept this as a consequence of holding this view. They have to be willing to impute some state of mind to hypothetical thinkers from a point in the past, even if these people are by definition not real.

10.3.5. Two Solutions for Retrocausality

I contend that the retrocausality problem must be caught on one of the horns of the following dilemma: either law exists at all times (at least for some purposes), or law is ontologically under-determined in a manner analogous to quantum states. Neither of these should be seen as attractive theoretical positions for a legal theory to accept, and so they reinforce the substantial difficulty presented by the retrocausality problem.

10.3.5.1. Resolution One: Eternal Law

The first potential resolution to the retrocausality problem is to embrace eternal law. Munzer has described this kind of view in explaining, and criticising, what he calls the 'changeless status theory' of retrospectivity. According to this theory:

> [A] given instance of performing [a] generic act is permissible only if there is (timelessly) no law prohibiting it. If a prohibitory retroactive law is passed, prior occasions of scratching one's nose are and have always been unlawful, though before the law was enacted it was impossible to know this—which is said to explain why such a law is considered morally objectionable. Act-instances forbidden by a retroactive law will be said always to have had that character in virtue of that law.[71]

Munzer objects to this theory, suggesting that it is artificial to suggest that the legal status of an instance of an act must be invariable, as this clashes with our intuition that retrospective laws actually *change* the legal status of past actions. It is true to say that there is some mismatch between this theory and our intuitions, but in a

[71] Stephen Munzer, 'Retroactive Law' (1977) 6 J Legal Stud 373, 377.

sense this objection begs the question against the void *ab initio* theory. It seems to assume that the void *ab initio* theory is wrong because it clashes with our intuitions, which in effect rejects the idea of objective law as false. This might be intuitive, but it is not a criticism that squarely addresses the void *ab initio* theory on its own terms.

The more significant objection Munzer points to, however, is that this changeless status theory is inapt to explain certain simple situations.[72] Consider an act that was at one time forbidden but is subsequently made permissible by a retrospective law. The changeless status theory now suggests (on the assumption there is no further reversion later in the legal system) that the act is timelessly permitted. But someone who was imprisoned on foot of that statute does not automatically walk free. Other legal devices, such as acquiescence, or doctrines of finality, can intercede and maintain the incarceration.[73] Given that this is so, Munzer claims that the changeless status theory here misdescribes what is going on. On the one hand, the act was never unlawful. On the other hand, the consequences of the conviction remain authoritative. From this, Munzer claims that it must be incorrect to describe the act as timelessly permitted.[74]

The account of 'creativity' accommodated by the eternal law response is also somewhat lacking. Recall that the difficulty of the retrocausality problem is simultaneously accommodating subjectivity, that is, judicial creativity, and the practice of finding unconstitutional laws void *ab initio*. The difficulty, as described in fig 10.1, is that this seems to require that legal decisions from future points in time can bear on what the law is *now*. At a minimum, this requires that the future is already 'real' in some sense.[75] If the future is already actual, then it follows that the account of creativity required by the eternal law answer effaces human agency in a way that some

[72] Munzer, 'Retroactive Law' (n 71) 378.

[73] Munzer treats this as hypothetical, but it will be recalled that this is in essence the exact problem the Irish Supreme Court faced in *A v Governor of Arbour Hill Prison* (n 1).

[74] Munzer, 'Retroactive Law' (n 71) 378.

[75] There is a school of thought in metaphysics that believes exactly this. It is usually known as 'eternalism': Theodore Sider, 'Presentism and Ontological Commitment' (1999) 96 Journal of Philosophy 325.

might find difficult to square with a robust concept of creativity. What is required is something like predetermination; judges can create law, but they can only create the law that they were bound or 'destined' to create, so to speak.

Finally, the eternal law answer requires not only that these future legal determinations exist (and so we are bound to make them at some point) but also that they can interact with laws that are (relative to them) in the past. This requires something like what is occasionally known as 'backwards causation'.[76] These difficult and unintuitive theoretical commitments are a heavy cost to bear to retain the view that unconstitutional law is void *ab initio*.

10.3.5.2. Resolution Two: 'Quantum Law'

The first resolution above assumed that there is, at any time, some truth to be found as to whether a legal instrument is constitutional or not. The eternal law argument is supposed to make sense of the fact that the truthmaker for this statement may be in the future, relative to where the observer in the legal system is situated, but it is extant, and it is completely determined. A different way to parse voidness *ab initio* and make sense of judicial rhetoric is to analogise with a concept from quantum mechanics: specifically, how quantum particles may be *superpositioned*. Superpositioning refers to the property of quantum mechanical objects that they may, for example, occupy two spaces at once. Schrödinger's cat is a well-known thought experiment in this regard.[77] The thought experiment runs as follows: a cat is placed in a box with some poison. The random decaying of a radioactive particle releases the poison. There is no way of knowing, without looking in the box, whether the cat is dead. According to the so-called Copenhagen interpretation of quantum mechanics, the cat is both alive and dead until observed, at which point the quantum state

[76] Michael Dummett, 'Bringing About the Past' (1964) 73 The Philosophical Review 338; Rebecca Roache, 'What is it Like to Affect the Past?' (2015) 34 Topoi 195.

[77] Erwin Schrödinger, 'Die gegenwärtige Situation in der Quantenmechanik' (1935) 23(48) Naturwissenschaften 807.

collapses into one actuality or another.[78]

If this sounds implausible, it is because it is.[79] Indeed, Schrödinger intended to posit his cat hypothetical as a *reductio ad absurdum* of the Copenhagen interpretation. Although the cat has ironically become a kind of mascot for the Copenhagen interpretation, the original connotations of absurdity are more appropriate here. On the face of it, law could not possibly be a quantum mechanical system in the way envisaged by the Schrödinger's cat example, as there is no relevant random subatomic event that may or may not occur to collapse the wave function of different probabilities into one actuality. It would be ridiculous to argue that law has quantum, subatomic properties. The quantum law argument cannot be dispatched quite so simply, however. Retrocausal practice does share an important feature with quantum systems: it is contingent on *observation* to ascertain what state a law is actually in. The quantum description might therefore function more as an analogy than a direct comparison.

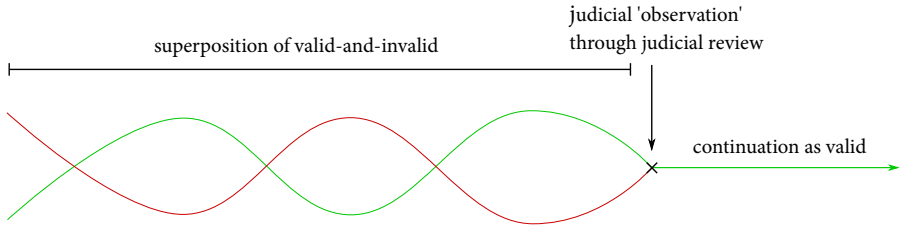
Applying this to the way the 'void *ab initio*' view parses constitutionality, one could argue that the constitutionality of a law is determined when judges review a legal instrument, but it is not thereby *changed* by the judge. It simply *is* in one state after the fact of observation. In this respect there is some analogy to be drawn between quantum superpositioning and one plausible interpretation of the void *ab initio* theory. It may be even more surprising still to see that philosophers of law have actually come much closer to this view than might be thought. Raz, for example, has opined that in cases where there are gaps in the law (say, where the law makes some moral norm a condition of legal validity) it is the case that: 'The pro-

[78] One problem with the thought experiment is that cats are not quantum mechanical objects. The system is meant to qualify as quantum mechanical because of its dependence on a quantum particle in a superposition (the radioactive particle that randomly decays). This is meant to translate into a more macroscopic quantum system involving the cat and the alive-and-dead superposition. In the present context I mean to draw comparison with quantum mechanics *allegorically* rather than directly, but it is worth nothing that this same criticism would apply to any kind of legal context.

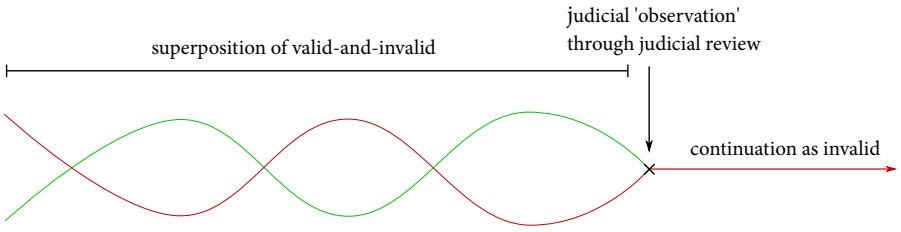
[79] Nevertheless, I am not the first person to use this type of analogy: Shivprasad Swaminathan, 'Schrödinger's Constitutional Cat: Limits of the High Court's Declaration of Unconstitutionality' (2013) 25 Nat'l L Sch India Rev 100.

position "it is legally conclusive that this contract is valid" is neither true nor false until a court authoritatively determines its validity'.[80] Similarly, Coleman once said that 'Philosophers generally agree that some sentences involving the application of vague predicates are neither true nor false'.[81] The statement that something 'is/is not the law' is one such predicate. These statements seem to impute more than mere epistemic uncertainty about the law (ie, statements as to our *knowledge* of what underdetermined law is bear no truth value until authoritative determination). They seem to go further and suggest that a statement as to what the law itself *is* has no binary truth value.

Consider the following diagram, illustrating how the superpositioning of legal validity would work:



Timeline 1: Superposition collapses into validity



Timeline 2: Superposition collapses into invalidity

Figure 10.2: Superposition Model of Legal Validity

There are, effectively, two realities in competition with one another: one in which

[80] Joseph Raz, 'Legal Reasons, Sources, and Gaps' in Joseph Raz, *The Authority of Law: Essays on Law and Morality* (2nd edn, OUP 2009) 75.

[81] Jules Coleman, 'Truth and Objectivity in Law' (1995) 1 Legal Theory 33, 49.

law-is-valid, and one in which law-is-invalid, and these collapse into one of the two upon judicial observation. Again, *qua* quantum theory there would be many problems with this model: why only judicial observation? If it was really a matter of quantum mechanics, there would not be any reason why judicial observation should be privileged over anyone else's observation. Again, it bears stressing that, the example is merely supposed to be drawing an analogy with quantum physics, and tendentiously at that. The point is that either the 'eternal law' model or the superposition model must be endorsed in order for the practice of voidness following legal invalidation to be effective.

It should be obvious that 'quantum' law presents no more attractive option to the legal theorist than did eternal law, and it should not be taken seriously as a commitment of legal theory.

10.3.6. Conclusion

The overarching goal of this section has been to show how the practice of finding unconstitutional law void *ab initio* is incompatible with a view of law as a subjective, dynamic concept. There is good reason to think that much judicial interpretation is, in fact, law-making,[82] and the subjectivity that this entails is difficult to reconcile with this common feature of constitutional practice.

I have suggested that these propositions may be reconciled by thinking that future legal norms both *already exist* and that they can interact with the legal norms of the present to make them unconstitutional. Alternatively, one may reconcile these propositions using the quantum superpositioning model. On the quantum model, one can accept that a law is both valid and invalid. This solution bites the underdetermination bullet by accepting what is, all things considered, an improbable ontology of law.

Neither of these responses is particularly attractive, given the deeply social and

[82] Scott Soames, 'What Vagueness and Indeterminacy Tell Us About Interpretation' in Andrei Marmor and Scott Soames (eds), *The Philosophical Foundations of Language in the Law* (OUP 2011).

pragmatic elements of law. Both require support of extreme ontological and metaphysical positions. Faced with such theoretical choices, the better choice may be to simply abandon one or more of our presumptions. The idea that judges have some creative function is an independently good one.[83] This leaves the view that voidness *ab initio* is the problematic presumption. I have already proposed that applicability is an alternative lens through which unconstitutionality may be analysed in chapter 8. The failure of the void *ab initio* model, and the concept of invalidity that it requires, to be accommodated under *either* an objective or subjective view of law is another reason to prefer applicability. I will return to further arguments in favour of applicability in chapter 11.

10.4. INSTITUTIONAL/MORAL PROBLEMS

This section considers objections that arise when the harshness of voidness *ab initio* is mollified either through the wholesale abandonment of that theory (prospective overruling), or the adoption of remedies that qualify the effect of the unconstitutionality (suspended declarations). Principally, they revolve around concerns that arise when unconstitutionality does not have immediate juridical effect, either for the individual litigant or generally. To be sure, these are serious difficulties, but they are not insurmountable, and they do not result in as difficult a challenge as the void *ab initio* theory faces.

Prospective overruling, as exemplified by *Linkletter*[84] in the United States and the post-*Damache*[85] case law in Ireland, faces the related difficulties of meting out individual justice in cases and motivating plaintiffs to take cases. It can also introduce an element of caprice in determining who should be able to avail of the benefit of unconstitutionality. At worst, litigants could be engaged in a race with one an-

[83] Many originalists in the United States might disagree with this view. However, outside the United States originalism is not as popular, and the (usually tacit) acceptance that judges must create law is less controversial.

[84] *Linkletter v Walker* 381 US 618 (1965).

[85] *Damache v Director of Public Prosecutions* [2012] IESC 11, [2012] 2 IR 266.

other to have a rule declared unconstitutional in their case first, as there may be a refusal to extend the benefit of the rule to other cases. This heavily individualistic attitude seems to be ill-at-ease with the public interest and common good objectives of constitutional rights litigation.

Suspended declarations face a different set of concerns. Recall that in this thesis I understand suspended declarations as preserving the applicability of legal norms that are deemed unconstitutional. The suspension thus freezes the applicability of the current law *in situ*, rather than deferring the invalidity to a later date. This raises similar plaintiff incentive problems as prospective overruling. It also faces rule of law difficulties in endorsing the continued and unqualified application of legal norms that a court positively knows to be constitutionally flawed. Because suspended declarations are viewed as a temporal measure, rather than one directed to applicability, the courts do not tend to allow the unconstitutional law to continue to have *limited* applicability in certain, suitable cases. Rather, the unconstitutional law is just given wholesale applicability, often for a period of a year if not more. This raises questions around whether blanket retentions of the applicability of unconstitutional law, even for limited time periods, can impede effective rights enforcement through judicial review.

Two particular concerns have been highlighted by Leckey with respect to the suspended declaration: (1) the position of the victim of the rights violation, and (2) the uncertainty and unpredictability as to whether the court will delay its order at all, as well as the reaction of other elements of the legal and justice systems thereafter.[86] These two issues are considered below.

10.4.1. General Justice at the Cost of Individual Injustice

This is the first of Leckey's concerns about the suspended declaration: blunting the immediate effect of unconstitutionality can leave the aggrieved and successful litigant without effective recourse. This difficulty applies equally to prospective overrul-

[86] Robert Leckey, 'The Harms of Remedial Discretion' (2016) 14 ICON 584, 591–96.

ing, and was particularly well-attended to by critics of the Warren court's prospectivity experiment.[87] Contrary to the supposed truism, *ubi ius, ibi remedium* ('where there is a right, there is a remedy'), it may transpire that the claimant has a right but no remedy. The legislative change that follows a court's suspended order may also fail to provide the victim with a remedy; there is no guarantee that the legislature will not amend the unconstitutionality with prospective effect only, for instance. Even if an individual remedy is granted to the applicant, the suspension may still worsen the position of those other than the litigant who suffer from the same infringement of rights, as it lengthens the duration of the rights infringement. Indeed, this assumes the legislature will take action at all: there is no guarantee the legislature will actually remedy the unconstitutionality within its allotted time.[88]

This concern becomes increasingly institutionally significant proportionately to the degree that plaintiff disincentivisation is endemic. This is because judges (at least in the systems under scrutiny in this thesis) mostly decide concrete and not abstract cases.[89] They need live issues before them to play the limited role in governance that the courts play; specifically, influencing policy to maintain constitutional standards, whether that is in the somewhat crude way envisaged by straightforward invalidation, or the dialogic model of the suspended declaration. If plaintiffs do not bring cases, the 'judicial' part in the 'judicial-legislative dialogue' never gets off the ground.[90]

This problem can be addressed through supplementary relief in the style of the South African interim order—an additional order to the declaration of unconstitutionality that gives more targeted relief to the plaintiff in the case. This would help

---

[87] Kermit Roosevelt, 'A Little Theory is a Dangerous Thing: The Myth of Adjudicative Retroactivity' (1999) 31 Conn L Rev 1075; Richard Fallon and Daniel Meltzer, 'New Law, Non-Retroactivity, and Constitutional Remedies' (1991) 104 Harv L Rev 1731.

[88] For example, this was the experience of the South African Constitutional Court in *Minister of Communications v Ngewu* [2013] ZACC 444, (2014) 3 BCLR 364.

[89] Recall the discussion of concrete and abstract judicial review in chapter 1.

[90] Leckey, 'The Harms of Remedial Discretion' (n 86) 596; Sujit Choudhry and Kent Roach, 'Putting the Past Behind Us? Prospective Judicial and Legislative Constitutional Remedies' (2003) 21 Supreme Court Law Review (2nd) 205, 247.

to mitigate many of the concerns raised here regarding suspended declarations, but it might be that for the advocate of legislative dialogue, the interim order would bring other difficulties to the table. For one, the requirement that judges establish some positive policy claims in the interim order might be taken as quasi-legislative, or tacitly suggesting a preferred resolution to the legislature.[91] Legislatures might be inclined to defer to the court's solution on constitutional issues, even if they are notionally being allowed free rein by the substantive declaration. This might defeat the point of the idea of judicial-legislative dialogue, depending on the degree to which the legislature becomes comfortable in conventionally adopting the courts' proposed solutions. This thesis does not take a position on the complex topic of dialogue and so this solution is at least *pro tanto* acceptable.

10.4.2. Response of Other State Institutions to Suspension

The suspended declaration, on the understanding in this thesis, asserts that the unconstitutional law is *invalid*, but that it will remain generally *applicable* for a determined period. This can leave participants in the legal system unsure how they should conduct themselves under this law. This was Leckey's second point, mentioned above. The suspension of the invalidity of the prostitution provision in *Bedford*,[92] for example, left the authorities in a confusing position with respect to enforcement of that law. Approaches to enforcement varied between provinces.[93] It seems difficult to justify continued prosecution (and punishment) under a provision tainted with unconstitutionality. Not only is there a moral conundrum in punishing someone under an unjust criminal law, but there are practical problems if those punished are likely to take appeals or collateral attacks because they feel their convictions are unfair. However, if this is the case it raises the question: what, if

[91] This is what Tushnet has called policy distortion: Mark Tushnet, 'Policy Distortion and Democratic Debilitation: Comparative Illumination of the Countermajoritarian Difficulty' (1995) 94 Mich L Rev 245.

[92] *Canada (Attorney General) v Bedford* 2013 SCC 72, [2013] 3 SCR 1101.

[93] Leckey, 'The Harms of Remedial Discretion' (n 86) 594, fn 64.

anything, is the effective difference between a suspended declaration of unconstitutionality and an immediate one if agents of the executive will not enforce the law once suspended? In such circumstances, is anything meaningful achieved through the act of suspension?

In many respects, this difficulty is a result of the way suspension of unconstitutionality is viewed as a *temporal* suspension of a determined, holistic effect of unconstitutionality (ie, invalidity). As I argued in chapter 8, this is very likely due to an absence of a theoretical and conceptual vocabulary that does not treat validity as the one controlling concept in unconstitutionality. Because validity is acontextual (a legal norm is either valid in the system or it is not), there is no scope for a court to say that it suspends the operation of the unconstitutional law in some cases but not others. Suspension here refers to holding the notional effects of validity at bay until a point in the future. However, on the basis of the distinction in chapter 8 between validity and applicability, there is an alternative way of understanding suspension. It could instead be viewed as *preserving* the applicability of the unconstitutional law (which is invalid by operation of a constitutional supremacy clause)[94] in a limited set of cases. This would allow a court to tailor relief in a way that might give subjects of the legal system more guidance as to how the unconstitutional law may or may not continue to have juridical and normative significance for them.

The effectiveness of this solution will be contingent, but they key point is that seeing suspension as an applicability measure allows for more finely-drawn relief than simply a blanket suspension of the *status quo*, including the identified constitutional faults of the impugned law. This coarse approach to suspension is more likely to lead to non-enforcement of the unconstitutional, but suspended, law. Where the court can give some guidance as to the circumstances in which the unconstitutional law should have continued application, it is more likely to avoid executive paralysis in the enforcement of that law.

[94] This is the case in both Canada and South Africa.

10.4.3. The Rule of Law

Concerns about the rule of law are bound up particularly with suspended declarations. In Canada, the suspended declaration awarded by the court in the *Re Manitoba Language Rights* case was specifically justified by reference to the rule of law, which the court also held to be a fundamental principle of the Canadian constitutional order reflected in the Constitution Acts of 1967 and 1982.[95] The Supreme Court offered a conception of the rule of law that comprised at least two basic features: (1) that law is supreme over all people and precludes exercise of arbitrary power, and (2) the creation and maintenance of an order of positive laws.[96] The former of these features is effectively a guarantee of separation of powers; the latter is a guarantee of a legal system. In the very unusual and extreme context of the *Manitoba Language Rights* reference, it is clear that the second of these rule of law principles would be violated. The complete eradication of positive law within the province of Manitoba is clearly something that the Supreme Court could not countenance while also maintaining itself as upholding the rule of law at the same time. Whatever one thinks of the current practice surrounding the grant of suspended declarations, if it was ever justified, it was in the *Manitoba Language Rights* reference.

Hogg has pointed out that the difficulty in *Re Manitoba Language Rights*[97] was that the two rule of law values were in contradiction:[98] the first requirement (the supremacy of law) was precisely what entailed that Manitoba's monolingual laws were invalid and void since they were contrary to the Constitution. But to hold fast to that value would necessitate a breach of the second principle, as it would leave Manitoba without a legal system. At the very least, then, *Manitoba Language Rights* could have stood for the more strict principle that where there is an outright rule of law contradiction within a legal system, that contradiction can be resolved by

[95] *Manitoba Language Rights* (n 1) [64].

[96] ibid [59]–[60].

[97] ibid.

[98] Peter Hogg, *Constitutional Law of Canada* (5th edn, Carswell 2007) 58-23.

ignoring or discounting certain other rule of law or constitutional principles. There is thus merit to Hogg's suggestion that the doctrine could have alternatively been justified on a doctrine of necessity in the face of constitutional crisis.[99]

There are other potential rule of law deficiencies in the practice of suspending declarations of unconstitutionality. Although the original justification of that remedy in *Manitoba Language Rights* and *Schachter*[100] was safeguarding the rule of law, this does not entail that in all cases a suspended declaration guarantees respect for that ideal. As Ryder has observed:

> Lawmakers might be getting the message that they take no significant risks if they pass laws without serious regard for Charter rights and freedoms. … The consequences of drafting laws that may violate the Charter, from a government's point of view, may be nothing worse than litigation and a second chance at drafting Charter-compliant legislation a few years down the road. In this way, a remedy initially designed to serve the rule of law now risks promoting its violation.[101]

This points to another important feature of the role of the courts in rights enforcement; not only does the court vindicate the rights of the victim, but also it implicitly reprimands the legislature for its failing. In classic cases like *Manitoba*, the rule of law failing was straightforward: 'there can be no rule of law without law'. Ryder's criticism, as quoted above, requires a more substantive picture of the separation of powers than the rule of law provides on its own. Nevertheless, it does suggest that the 'dialogue' between the legislature and courts need not always be one of deference of the latter to the former; occasionally, the courts will need to take an assertive stance with rights issues, and this is a stance ill-served by the suspended declaration where it becomes the default response to unconstitutionality.[102]

The distinction between these judicial attitudes has been succinctly captured by Leckey, who suggests that there are two judicial postures reflected in straight-

---

[99] Peter Hogg, 'Necessity in a Constitutional Crisis' (1985) 15 Monash U L Rev 253.

[100] *Schachter v Canada* [1992] 2 SCR 679 (SCC) 715–16.

[101] Bruce Ryder, 'Suspending the Charter' (2003) 21 Supreme Court Law Review 267, 288.

[102] For a less sanguine view of dialogue language, see: Aileen Kavanagh, 'The Lure and the Limits of Dialogue' (2016) 66 UTLJ 83. For a critical view on the theory in general, see: Eoin Carolan, 'Dialogue Isn't Working: The Case for Collaboration as Model of Legislative–Judicial Relations' (2016) 36 Legal Studies 209.

forward invalidation and the suspended declaration. The first is a constitutional enforcement posture, which treats constitutionalism, rights enforcement, and the rule of law as trumps over other concerns. The second is a legislative engagement posture, which treats the judicial role as mediating engagement between the Constitution and other limbs of government.[103] He goes on to acknowledge the connection between the second posture and weak-form judicial review,[104] leaving to implication the already apparent association between the first posture and strong-form review. In a similar vein, Roach has suggested a 'two-track' approach to constitutional remedies.[105] The analysis in this thesis helps to see how these increasingly recognised divisions track the distinction between applicability-focused remedies and validity-focused remedies. This scholarship on remedies is an important reminder to identify how these different approaches can be tailored to achieve different goals in constitutional law, and how the findings of this thesis, particularly the distinction between validity and applicability, can help to inform and reinforce such innovative proposals. This interaction with constitutional remedies is considered further in chapter 11.

10.5. CONCLUSION

This chapter considered three criticisms targeting aspects of practice on unconstitutionality. The first of these problems, the 'character of law problem', criticised 'right answer' theories of law for subscribing to an implausible account of objectivity in law. Right answer theories, as seen in chapter 8, map primarily to the legislative nullity answer to the derivation question. There is therefore good reason to prefer the judicial intervention answer, all other things held equal. The practice of voidness *ab initio* relies quite strongly on constitutional law being characterised as objective,

[103] Robert Leckey, 'Enforcing Laws That Infringe Rights' [2016] PL 206, 210.

[104] ibid 211.

[105] Kent Roach, 'Remedies for Laws that Violate Human Rights' in John Bell and others (eds), *Public Law Adjudication in Common Law Systems* (OUP 2015); Kent Roach, 'Polycentricity and Queue Jumping in Public Law Remedies: A Two-Track Response' (2016) 66 UTLJ 3.

and this criticism suggests that this characterisation fails.

The next criticism, the 'retrocausality' problem, focused on criticising an alternative subjectivist reading of voidness *ab initio*. That is, one in which the judge exercises a creative function. I argued that voidness *ab initio* as a consequence of unconstitutionality cannot easily be squared with the view that judges make law. This creative view of the judicial power is well-motivated. I argued that to rationalise the practice of unconstitutionality using a validity analysis, while simultaneously accepting that judges have some law-making function, requires the adoption of an implausible view that future law already exists, and that it can ground the unconstitutionality of law that pre-dates it. So, although voidness *ab initio* and judicial law-making can be reconciled, it comes at a high price. In combination with the character of law problem, this forms a strong basis for discarding the void *ab initio* analysis of unconstitutionality.

The final set of criticisms, the 'institutional/moral' problems were directed primarily at innovations in response to the strict void *ab initio* view. The most significant innovations in this regard are the suspended declaration of unconstitutionality, and the technique of prospective overruling. The problems here are not as theoretically severe as the character of law or retrocausality problems. However, they do raise concerns regarding the costs of prospectivity or retaining the applicability of unconstitutional law. These criticisms demonstrate that tampering with the effect of unconstitutionality is a difficult and context-sensitive question. They suggest that courts and legal practitioners should be aware of the advantages and disadvantages of retrospectivity or prospectivity on a case-by-case basis. I return to the idea that the imposition of retrospectivity should be assessed on more straightforwardly moral grounds in chapter 11. Perhaps most importantly, these criticisms suggest that the suspended declaration of unconstitutionality is not a panacea. Although it may be a useful remedial tool, understanding it through a more theoretically nuanced lens may help to overcome some of the difficulties with it that have been identified. Most importantly, understanding it as a modification to the effect of unconstitutionality (targeting applicability rather than validity) rather than temporal duration

(whatever the effect is, it obtains in the future and not now) brings significant clarity to the potential expansion and limits of the remedy.

Part IV.

Conclusion

11 Implications for Constitutional Theory and Practice

CHAPTER OVERVIEW

11.1. Introduction 325

11.2. The Preferred Theory of Unconstitutionality 327

11.2.1. Judicial Power as the Ground of Unconstitutionality 327

11.2.2. Inapplicability as the Effect of Unconstitutionality 328

11.2.3. Moral Considerations and the Temporality of Unconstitutionality 331

11.3. The Insufficiency of a Remedies-Centred Analysis 334

11.3.1. The Pragmatic Approach 335

11.3.2. The Conceptual Approach 337

11.3.3. Understanding the Suspended Declaration 340

11.3.4. Roach's Two-Track Remedies Proposal 341

11.4. Unconstitutionality and the Comparative Method 343

11.4.1. The Use of Foreign Law 343

11.4.2. The Limits of Textual Constraints 345

11.5. Conclusion 346

11.1. INTRODUCTION

In this closing chapter, I argue that the findings of the foregoing chapters can be used to formulate a new model for dealing with unconstitutional legal norms. As argued in chapter 8, there are three general questions a theory of unconstitutionality must answer, which I have termed the 'derivation' question, the 'effects' question,

and the 'temporal' question. In chapter 10, I argued that there are significant flaws with the most common theory of unconstitutionality: the void *ab initio* theory. This chapter seeks to supplement that critical work by suggesting a preferable model of unconstitutionality. To be sure, there is no 'silver bullet' rule of sufficient generality to deliver justice in all cases. One of the most important conclusions from the literature on public law remedies is that allowing for flexibility of approach is essential. The aim here is therefore more modest: I attempt to outline the model of unconstitutionality that is optimal in most cases, while allowing that it may not be followed in every circumstance. This is particularly so regarding the temporal questions. Unlike the derivation and effects questions, these are not really technical theoretical points *per se*; they have a much closer relationship to broader moral and institutional questions in adjudication. This means that a general rule may be a useful heuristic for these issues, but freedom to deviate from that rule when appropriate is vital.

The chapter offers some additional remarks on the shortcomings of an exclusively remedial analysis of unconstitutionality. Many public law commentators prefer to cast unconstitutionality as a subset of problems of judicial remedies, institutional design, or constitutionalism generally. There is substantial merit in understanding how unconstitutionality intersects with all these problems. However, the failure to consider unconstitutionality as a standalone topic of theoretical interest has impoverished discourse on the subject. The theoretical reflections of this thesis make a valuable contribution in attending to this deficit and thus adding further improvement and insight to remedial discourse.

This chapter also notes some implications of this thesis for comparative constitutional law more generally. In particular, it considers the use of foreign law as persuasive precedent in cases considering the nature and effects of unconstitutionality. It has been argued by some authors recently that if there is a paradigm context for the use of foreign law by judges in constitutional cases, it will be around issues such as unconstitutionality. This chapter demurs from this view, using the confusion that has arisen in India over the nature and effects of unconstitutionality as an example. Where points of theoretical nuance can be attended by the judiciary then

the use of foreign law may be effective in this area, but as judges often do not have the luxury to consider such complex points in detail, it may lead to more confusion than clarity.

11.2. THE PREFERRED THEORY OF UNCONSTITUTIONALITY

In this section I present the most preferable answers to each of the derivation, effects, and timing questions asked in chapter 8. 'Most preferable', in this context, does not mean that these answers respond perfectly to every situation where unconstitutionality arises. Rather, these answers represent what the better view on balance is, taking into account that some views offer greater flexibility than others.

11.2.1. Judicial Power as the Ground of Unconstitutionality

The legislative nullity view can produce significant difficulty in practice. As I argued in chapter 8, this view usually must be accompanied by some commitment to the origin of unconstitutionality being rooted in the past (see Table 8.1 for a summary). One of the more putatively attractive motivations for adopting the legislative nullity view is its ability to produce retrospective effect without actually making the revised constitutional norm retrospective in the true sense of the term.[1] This is because the revised constitutional norm, whatever it is, was always present (and the judiciary merely notice it at some later point). This allows the norm to achieve retrospective effect without being effective for a span of time that pre-dates its own coming into existence. It is instead effective for its entire existence, which happens to start from the point in the past when the Constitution first came into effect.

The judicial intervention theory is a much more plausible account of the generation of unconstitutionality. This view simply admits that the judiciary play a constitutive role in *making it the case* that some norms are unconstitutional. This is, to be sure, a significant power to afford the judiciary but, as I outlined in chapter

[1] That is: changing the present or future legal status of actions or events that occurred *before the law came into effect*. See further: Stephen Munzer, 'Retroactive Law' (1977) 6 J Legal Stud 373, 381; Charles Sampford, *Retrospectivity and the Rule of Law* (OUP 2006) 22

1, this thesis takes strong-form judicial review for granted. If strong-form review is assumed, I argue that giving the judiciary the power to regulate constitutionality directly in this way is inevitable. It is no real constraint on the judicial power to require the presentation of unconstitutionality as a constitutional *fiat*, as the legislative nullity view does. All this achieves is an alternative, and more obfuscated, semantic presentation of the same operation and result.

Another advantage of the judicial intervention theory is that it is less restrictive in terms of the answers to other questions around unconstitutionality that it can pair with. It has no tension with any of the other answers and so allows for 8 possible combinations. This is more flexible than the legislative nullity view, which has only 4 combinations. The 'default' position in Ireland, South Africa and Canada maps on to variants of the legislative nullity view. It is no accident that the suspended declaration has arisen in each of these jurisdictions, as this is an innovation that can preserve a commitment to the legislative nullity view while avoiding some of its more undesirable implications.

11.2.2. Inapplicability as the Effect of Unconstitutionality

This thesis recommends that invalidity should not be treated, as it often is, as the focal point of analysis in unconstitutionality. The United States is a model of best practice here, with several judicial statements overtly favouring the view that the federal judiciary has no power to regulate the validity of statutes. In other jurisdictions the constitutional text fetters practice to a greater degree, with references to validity specifically in the Irish and South African Constitutions, and a reference to void(ness) in the Indian Constitution.

An alternative concept, identified in chapters 8 and 9, to replace the validity analysis is the concept of *applicability*.[2] It will be recalled that while validity is a question of the genealogy and content of legal norms (who made them, and are

[2] Interestingly, this would be the most natural reading of the Canadian Constitution's language of 'no force or effect'. The Canadian courts have, to an extent, foreclosed on this possibility by treating the language as requiring unconstitutional law to be invalid and void.

they consistent with other norms of the system?) applicability is a question of the scope and justificatory power of a legal norm (what cases does it cover, and is there a norm of the system that directs a judge to apply it in such cases?).[3] The norms that might direct a judge to apply a legal norm might be general and parasitic on validity; for example, 'all valid, non-repealed statutory norms that cover a case are to be applied to generate a result in that case' is a plausible applicability norm. For this reason, validity and internal applicability will usually coincide. Navarro and Rodríguez have made this same observation (recall that internal applicability refers to the specific case falling within the generic type of case covered by a norm):

> [I]nternal applicability and membership in the legal system jointly count as an ordinarily sufficient condition for a certain norm to belong to the set of applicable norms with respect to case *c*. In this sense, judges cannot arbitrarily ignore those valid norms regulating a certain case; and if it is the judge's ultimate decision that they be disregarded, she has to offer a justification for so doing. What surely cannot be accepted is that internal applicability and membership in the legal system are sufficient to warrant that a certain norm will defeat all other competing applicable norms regarding case *c*.[4]

However, vested rights and legitimate expectations can work to generate applicability norms that are not parasitic on validity in this way. Often the threat of injustice that arises in cases of unconstitutionality is the systemic shock that disrupts a settled state of affairs. The concepts of acquiescence and equity that judges frequently cite to in justifications of retrospective effect seem to tacitly indicate a conventional applicability norm that can be used to apply unconstitutional law.

One of the principal advantages of an applicability analysis is that it is *relative to individual cases*. In some respects, the difficulties generated by findings of unconstitutionality are caused by the *erga omnes* effect of declarations. As a result of unconstitutionality with this effect, the juridical position of all the citizens of a jurisdiction can radically shift. This is because validity is, as Bulygin had it, a four-place relation between a legal system, an empowering norm of this system, an act of norm

---

[3] Recall that, as discussed in chapter 8, these bracketed questions refer to internal and external applicability, respectively.

[4] Pablo Navarro and Jorge Rodríguez, *Deontic Logic and Legal Systems* (Cambridge University Press 2014) 138.

creation, and a norm created by means of this act.[5] Applicability, by contrast, is a five-place relation between a legal system, an application norm that belongs to this system, a court, a case, and a norm that is to be applied.[6]

Applicability, therefore, is always relative to a case or, at broadest a *type* of case, that is unified by a particular feature. It also does not have 'an act of norm creation' as one of its *relata*. This gets around the awkward issue of determining whether a law 'exists' (is a member of some system) or not, for the purposes of an applicability analysis. All that matters is the legal system, a member application norm, a court, a case, and a norm to be applied. This allows a court to avoid making grave statements about the existence or non-existence of a law, and it more naturally lends itself to a case-specific analysis.

Having judges regulate applicability rather than validity may also be preferable from a separation of powers point of view. The legislature has the capacity to issue valid legal norms. If the judiciary have a power to unwind or destroy these norms, it could make the legislature's task very difficult. One problem is that on the voidness *ab initio* theory that so often accompanies a finding of invalidity, the legislature cannot amend the stricken law. This is because there is, strictly speaking, nothing to amend. Instead, the legislature has to replicate the entire section. It would also presumably have to explicitly give this section retrospective effect if it were to govern incidents that occurred before it was passed.

Applicability is more flexible than validity when it comes to revising unconstitutional law because it is naturally more case-specific, and it can be taken to be a property of law that is specially regulated by the judiciary. A judge is relatively free to find legislation inapplicable to one set of facts and applicable to another. A later judge might reverse these findings and make their new judgment retrospective. None of this requires the engagement of slow-moving and deliberative legislative machinery. It is also not as capricious as it might seem, as precedent will bind judges and make

[5] Eugenio Bulygin, 'The Problem of Legal Validity in Kelsen's Pure Theory of Law' in Carlos Bernal and others (eds), *Essays in Legal Philosophy* (OUP 2015) 321. See chapter 8.

[6] ibid 321.

them slow to revisit and modify understandings of whether a law is applicable to a particular type of case. Finally, this does not impede cooperation between the legislature and judiciary. Indeed, it is likely easier for these powers to coordinate where they are not interfering with one another on issues of legal validity. If the legislature does not approve of a court finding on applicability, it can render that law invalid and replace it with a different one. Although I argued that invalidity does not imply inapplicability, it is hard to imagine a judge preferring invalid law where valid law exists and has been specifically provided by the legislature to address the same types of case. The mere fact that invalidity does not inexorably *require* inapplicability does not itself entail that there will not be other good reasons to apply valid law.

A difficulty in using applicability is that some constitutional texts expressly require a validity analysis (see chapter 2). However, in this chapter I also noted that each of the Constitutions studied in this thesis has some kind of assent procedure, which I described under the heading 'Constitutional Conditions for a Bill to Count as Law'. It is possible to interpret these provisions as endorsing a distinction between validity and applicability. The legislature can pass a legally valid Bill (which can become a legally valid Act), but it is only after passing the assent procedure that this becomes applicable or enforceable law. Although assent procedures in most countries may seem to be mere formalities, the possibility of reading this distinction into constitutional texts may help to circumvent what might otherwise seem to be an inescapable commitment to a validity analysis. To be sure, further detailed work would need to be done in each of the five jurisdictions to fully substantiate this claim, but the analysis in this thesis at least opens the possibility of this line of argument. Such arguments are worth considering, given the general merits of an applicability analysis.

11.2.3. Moral Considerations and the Temporality of Unconstitutionality

The void *ab initio* theory does not require much thought to be given to the temporality of unconstitutionality. Because it insists on the unconstitutional law being treated as a nullity, the automatic standard will be that unconstitutionality is fully

retrospective. An alternative theory will not have the benefit of this somewhat fictitious way of presenting the temporal issues. It must therefore supply an alternative analysis of the justification of retrospective law.

Given the theory of legal systems that I supported in chapter 9, there is good reason to treat unconstitutionality adjudication as effectively requiring a judge to choose between sets of applicable law in various momentary legal systems. Some of these sets may contain invalid norms. I expand on this further below in the discussion on remedies, using an 'intertemporal conflict of laws' metaphor. For present purposes, what is important to note is that the underlying justification for conflict of laws rules—doing fairness to the parties[7]—is also a fundamental concern in cases involving the applicability of unconstitutional law. This gives further animation to this metaphor.

One of the core concepts in conflict of laws/international private law that helps to realise the goal of doing fairness to the parties is the idea of 'connecting factors'.[8] These are numerous factors of a case that might connect one or more of the parties to a particular territory for the purpose of the choice of law.[9] These connecting factors could be viewed as application norms governing membership of the set of norms applicable to the case, as discussed in chapter 9. Of course, for the question as it arises in private international law, all of these criteria are necessarily spatial. Given that private international law is concerned with issues of the proper forum and law to apply to multi-territorial disputes, this makes good sense.

To extend the metaphor more meaningfully to multi-*temporal* disputes, different connecting and disconnecting factors will need to be chosen. There is already

---

[7] Max Rheinstein, 'Place of Wrong: A Study in the Method of Case Law' (1944) 19 Tulane L Rev 4; *Tolofson v Jensen* 1994 CanLII 44 (SCC), [1994] 3 SCR 1022, 1058; Adrian Briggs and others (eds), *Dicey, Morris and Collins on The Conflict of Laws* (15th edn, Thomson Reuters 2012) [1-005].

[8] Adrian Briggs and others (eds), *Dicey, Morris and Collins on The Conflict of Laws* (15th edn, Thomson Reuters 2012) [1-079].

[9] A variety of connecting factors are used in different areas of the conflict of laws: domicile (*lex domicilii*); residence; nationality (*lex patriae*); place where a thing is located (*lex loci situs*); place where a tort is committed (*lex loci delicti*); place where a contract is made (*lex loci contractus*); place where a contract is to be performed or a debt is to be paid (*lex loci solutionis*); place where a marriage is celebrated (*lex loci celebrationis*).

some analysis of this in the context of retrospectivity; retrospectivity doctrines in law ask when, if ever, it is appropriate to apply the law as it is now to situations that occurred in the past. Retrospectivity is already about choosing between the law of the past legal system and the law of the present legal system.

Sampford's work on retrospectivity is instructive in this regard; he points to four different justifications that may operate in favour of retrospectivity:[10]

(1) the 'better rule' justification: where a rule replaces an older rule, and the new rule is all things considered better than the old rule, then the new rule should be given as wide an ambit as possible (though this may need to be weighed against reliance interests in the old rule);

(2) the 'better authority' justification: it may be the case that the new rule is more democratically legitimate than the old rule, particularly if it enjoyed tacit support before it was enacted (ie, if the change was adumbrated and was popular before enactment);

(3) the 'efficiency' justification:[11] retrospective law may be both easier to apply (eliminating a distinction between the 'law then' and the 'law now') and more effective (transition periods between legal regimes give people who stand to lose benefits under the new scheme to prepare and even attempt to frustrate the effort of reform); and

(4) the 'fairness' justification: some reliance interests are not worth outweighing if they are otherwise iniquitous.

Each of these justifications is only *pro tanto*; it might be outweighed by other moral considerations in any given case. But the key point that Sampford establishes is that there are genuine moral reasons why retrospectivity might be justified. This

[10] Charles Sampford, *Retrospectivity and the Rule of Law* (OUP 2006) 229–42.

[11] This justification has gained particular support from certain law and economics theorists. Sampford cites to: Michael Graetz, 'Retroactivity Revisited' (1985) 98 Harv L Rev 1820; Louis Kaplow, 'An Economic Analysis of Legal Transactions' (1986) 99 Harv L Rev 509.

runs contrary to the often-assumed view that retrospectivity is never really truly morally justified, but it may be used where it is the least unjustified of several bad options. These justifications form a kind of 'connecting factor' analysis, connecting the case either to the law as it was at some time in the past, or to the law as it stands on the day of final hearing in a trial.

These justifications will then compete with other interests that operate as sort of 'disconnecting factors'—efficiency, reliance, fairness, etc—where a question of retrospectivity arises. Sampford summaries how these factors might contend with each other:

> In general, the fairness of retrospectivity will increase or decrease in inverse proportion to the reasonableness of the reliance involved: the more reasonable the reliance, the less fair it is to fail to protect that reliance interest from retrospective effect. Conversely, where reliance is extremely *un*reasonable, fairness will come out in favour of retrospective application.[12]

It might come as something of a disappointment that no bright line rule emerges from this analysis. However, any such rule in the context of retrospectivity and unconstitutionality would be a red herring. No more than in the conflict of laws, the 'rules' for choice of law can merely operate to draw the judge's mind to the relevant considerations. How those are weighted will invariably be determined on a case-by-case basis. This transparency is, however, itself a virtue. It is better to foreground these values so that judges and advocates can consider and contest the real factors that do and should influence the ultimate decision in a case. The approach to unconstitutionality endorsed in this thesis helps to achieve this foregrounding.

11.3. THE INSUFFICIENCY OF REMEDIES-CENTRED ANALYSIS

There are at least two ways of styling unconstitutionality as a legal problem, which I call the 'pragmatic' and 'conceptual' approaches here. Broadly speaking, the pragmatic approach casts the issues around the declaration of unconstitutionality as a

[12] Sampford (n 10) 247–48. Emphasis original.

problem of remedial discretion within the broader public law context. The conceptual approach, on the other hand, styles unconstitutionality as a specific instantiation of a more general conceptual problem that might be called the 'intertemporal conflict of laws', as already alluded to above.

These are not competing and mutually exclusive alternatives, but the pragmatic analysis is often favoured over the conceptual one. I argue below that although there are highly important practical considerations that arise in difficult cases of unconstitutionality, this should not remove the conceptual underpinnings of unconstitutionality from consideration. A greater appreciation of these more nuanced conceptual points can help remedial flexibility to be exercised in a more predictable and principled way. A failure to appreciate these points risks a remedies-centred analysis being abstracted from the particular context of declarations of unconstitutionality and being cast more generically as an institutional or separation of powers concern. I argue that the conceptual approach can usefully supplement and buttress the pragmatic approach to generate a more satisfying and enlightening understanding of remedies for unconstitutionality.

11.3.1. The Pragmatic Approach

The problem of unconstitutional statutes is sometimes understood through the lens of constitutional remedies rather than as an aspect of constitutional or legal theory.[13] Rather than being a function of some property of law or other theoretical commitment, the response to unconstitutionality is seen as a remedial choice on the part of courts. This position has been advanced by commentators in South Africa[14] and a similar proposition is advanced by Craig concerning English administrative

[13] Richard Fallon and Daniel Meltzer, 'New Law, Non-Retroactivity, and Constitutional Remedies' (1991) 104 Harv L Rev 1731; David Kenny, 'Grounding Constitutional Remedies in Reality: The Case for as-Applied Constitutional Challenges in Ireland' (2014) 37 DULJ 53.

[14] '[A]s a practical matter, inconsistency, invalidity and remedies cannot be separated from one another. … Invalidity … follows from inconsistency with the constitution but, by *declaring* the law or conduct to be invalid, a court grants a remedy'. Iain Currie and Johan de Wall (eds), *The Bill of Rights Handbook* (5th edn, Juta 2005) 193.

law noting that while the fundamental position should be that invalid acts are retrospectively void, the courts have a discretion to grant or deny a remedy in response to such acts irrespective of their (in)validity.[15] The focus on judicial remedies may help to ameliorate some immediately practical concerns with this area of law, but as Feldman observes: '[f]ocusing on judicial discretion separates the juristic status of a decision from its legal consequence. On this view, "voidness" is merely a threshold condition for the grant of a remedy'.[16]

That voidness should not be made a function of remedies is illustrated in the somewhat laboured reasoning in the following quotation from an older edition of Wade's administrative law textbook:

> Although action which is adjudged to be *ultra vires* is properly described as void or a nullity, this voidness necessarily depends upon the right remedy being sought successfully by the right person. … For as against third parties, whose rights are not infringed, a 'void' act may well be valid if they have no legal title to challenge it. Even the injured party may be refused relief, e.g. by an exercise of discretion or because of some waiver. The meaning of 'void' is thus relative rather than absolute; and the court may in effect turn void acts into valid ones by refusing to grant remedies. There is no absurdity in this. The absurdity lies rather in supposing that 'void' has an absolute meaning independently of the court's willingness to intervene.[17]

It is difficult to start from the premise that an action is 'properly described' as void or null, and then move to ameliorate this by making voidness 'relative'. This is precisely the mistake that mired the Indian courts in difficulty.[18] In this case, voidness seems to be relative to the remedy sought. But voidness is synonymous with non-existence, and existence is absolute, not relative; it seems difficult to imagine how something can not exist in one context and exist in another. A proper concept of 'relative voidness' thus has a lot more work to do than simply being a remedial *fiat*. Another issue with the reasoning in the excerpt from Wade is that it assumes that the court 'turn[s]' void acts into valid ones by refusing to grant remedies. In other words,

---

[15] Paul Craig, *Administrative Law* (7th edn, Sweet & Maxwell 2012) ch 24.

[16] David Feldman, 'Error of Law and Flawed Administrative Acts' (2014) 73 CLJ 275, 281.

[17] Henry Wade, *Administrative Law* (3rd edn, Clarendon Press 1971) 105–6. Cited in: Feldman (n 16) 280–81.

[18] See generally chapter 6.

it is the remedy that causes or brings about the voidness. This is often not what courts consider themselves as doing, as this type of act would require something like a negative legislative power, or a power of repeal. Such a power is not usually thought to properly rest with the judiciary.

It is important to recall that a finding of voidness is a theoretical artefact, and rather than curing the symptom (through remedies) it might be better to seek to cure the disease (the finding of voidness itself). The English perspective on invalidity and voidness in administrative law is preoccupied with the symptom, and it has thus produced a confusing literature and unhelpful conceptual distinctions such as the idea of 'relative' voidness, which plagued the Indian courts,[19] and distinctions between 'legal' and 'factual' validity, which was recently (and unfortunately) endorsed by the South African Constitutional Court.[20] However we resolve the ontological and moral problems that surround the invalidation and inapplicability of laws, courts need to be able to address practical problems on both an individual and institutional level. However, rather than continuing to rely on *ad hoc* supplementary doctrines to temper these undesirable effects of unconstitutionality, my proposal is that revising those core theoretical commitments that cause the difficulty in the first place is a more sustainable and long-term desirable cure than continuing palliative avoidance through limiting doctrines.

11.3.2. The Conceptual Approach

Contrary to the above remedial cast, some judges occasionally speak of unconstitutionality (particularly its implications for retrospectivity) as being an issue of the 'intertemporal conflict of laws'.[21] The temporal effect of judicial decisions is an un-

[19] *Behram Khurshed Pesikaka v State of Bombay* [1954] INSC 15, [1955] 1 SCR 613.

[20] *Oudekraal Estates (Pty) Ltd v City of Cape Town* [2004] ZASCA 48, (2004) 3 All SA 1 [29]–[30].

[21] See, for example: *James B Beam Distilling v Georgia* 501 US 529, 535 (1991). See also Judge Harlan's dissent in *Mackey v United States* 401 US 667, 682 (1971), where he discussed the problem of the 'choice of law' to apply to *habeas* petitioners.

derstudied topic in both comparative law and legal theory.[22] This is surprising, considering the relative ubiquity of this aspect of law. It is common sense that judicial decisions, and the conduct that they regulate, occur in time. However, there may be changes in the law that change sanctioned conduct into permissible conduct. Unlike the normative guidance we may receive from moral reasoning, the guidance we receive from law is time-dependent. This also means that the legality of an agent's conduct is time-dependent. As Higgins has observed: that the question of 'whether a court today finds the law to be X, is the State liable for all past occasions in which it has applied it as if it were Y' is a 'problem common to all systems of law'.[23]

This observation invigorates the question of an intertemporal 'conflict of laws'.[24] While conflict of laws is usually thought of in a territorial or jurisdictional way, it is fundamentally about rules that govern disputes where two different legal systems may be applied to a dispute. When one considers that the legal system is bounded by time as well as by place, this can be conceptually understood as two legal systems (a past one and a present one) vying over the dispute. This is easily captured by recalling from chapter 9 that a legal order (or non-momentary legal system) will be comprised of many momentary legal systems.

Intertemporal conflicts problems become particularly acute in systems that possess both a written constitution and strong-form judicial review, for the reason that written constitutions serve to make clear what legal guidelines the constitution sets, and strong-form judicial review brings into more acute focus conflicts between

---

[22] Though there are signs of this changing now: Patricia Popelier, Sarah Verstraelen and Dirk Vanheule (eds), *The Effects of Judicial Decisions in Time* (Intersentia 2013); Eva Steiner (ed), *Comparing the Prospective Effect of Judicial Rulings Across Jurisdictions* (Springer 2015).

[23] Rosalyn Higgins, 'Time and the Law: International Perspectives on an Old Problem' (1997) 46 ICLQ 501, 507.

[24] John McNulty, 'Corporations and the Intertemporal Conflict of Laws' (1967) 55 Ca L Rev 12; Julie O'Sullivan, '*United States v Johnson*: Reformulating the Retroactivity Doctrine' (1983) 69 Cornell L Rev 166, 170; John Randall Trahan, 'Time for a Change: A Call to Reform Louisiana's Intertemporal Conflicts Law (Law of Retroactivity of Laws)' (1999) 59 La L Rev 661; Jackie McCreary, 'Retroactivity of Laws: An Illustration of Intertemporal Conflicts Law Issues through the Revised Civil Code Articles on Disinherison' (2002) 62 La L Rev 1321; Ciarán Lawlor, 'Troubling Times: Intertemporal Law and Theories of Approach to the Effects of Unconstitutionality' (2008) 7 COLR 71.

these constitutional norms and legislative norms. Constitutional law has the distinguishing feature of being an aspect of public law, and thus having greater potential to achieve *erga omnes* effect; it applies to anyone and everyone, and it is often framed in such general terms that it bears on many peoples' normative position.

On the basis of the theory of legal systems favoured in chapter 9, the intertemporal conflicts metaphor seems apt. On that theory, unconstitutionality describes a situation where there are several momentary legal systems in competition, at least one of which includes the unconstitutional norm in its set of norms applicable to the case. A judge must then choose which of these momentary systems properly governs the case. The appropriate system to govern the case might be chosen for several reasons, which recall the 'temporal connecting factors' discussed in the context of Sampford's work, above. One momentary system might have been the one that existed at the time the relevant facts occurred. A different momentary system might contain a more just or fair rule. Importantly, validity does not control the analysis here. A rule might be invalid *because* it lacked in justice or fairness, but if it is not applicable it is because of these reasons of justice and fairness themselves, not because of some automatic operation of invalidity. Choice between momentary legal systems is therefore a question to be resolved on moral criteria such as legitimate expectations, weight of harm, justice of the rule, etc.

Thus, I suggest that conceptualising unconstitutionality more along the lines of conflict of laws has some merit. It focuses on unconstitutionality as a specific phenomenon and situates it in a detailed and compelling theory of legal systems. Remedial analysis naturally focuses disproportionately on the temporal questions of unconstitutionality, but it tends to take the effect of unconstitutionality for granted and is unable to rationalise how unconstitutional norms can still have significance in the legal system. This is not to say that remedial analysis is inappropriate; it is merely *insufficient* by itself. If it is used to supplement, rather than replace, a considered theory of unconstitutionality then greater clarity and nuance could be achieved in this area of practice. The theory of unconstitutionality advanced in this thesis has attempted to bridge this gap between the pragmatic and conceptual approaches.

11.3.3. Understanding the Suspended Declaration

The suspended declaration of unconstitutionality presents a useful case-in-point of how greater theoretical understanding can enhance our understanding and application of remedies for unconstitutionality. As I have argued in chapters 8 and 10, the understanding of the suspended declaration as just being a modified temporal gloss on the fundamental theory of voidness *ab initio* is misguided. This temporal understanding ignores a serious contradiction between two ideas. On the one hand, voidness *ab initio* is produced from a legislative nullity, generated by autonomous operation of the constitution as supreme law. On the other hand, the idea that the judiciary can *suspend* unconstitutionality through the grant of an order suggests that it is the judicary's order that 'constitutes the unconstitutionality'. In other words, to view the suspended declaration as a temporal doctrine, one needs to simultaneously assert both that the judiciary do not have a constitutive role in making unconstitutionality the case, and that they do have such a role. This is a plain contradiction.

The preferable analysis is to consider 'suspension' not as referring to deferring or delaying the invalidity, but as prolonging the applicability of the unconstitutional law in spite of that invalidity. This accurately describes what this remedy achieves, and eschews temporal language that does more to restrict the scope of the remedy rather than widen it. Consider that many of the difficulties identified with the suspended declaration in chapter 10 concerned the potential that a suspended declaration would not afford an individual an effective remedy. The applicability understanding of the suspended declaration can help to respond to some of these challenges, as it lends legitimacy to preserving the effect of the unconstitutional law only as it *applies* to certain cases. The belief that suspended declarations modify validity impedes this sort of modification, as it is not possible to declare a law valid in some circumstances but invalid in others. Applicability analysis can therefore add both theoretical legitimacy and greater flexibility to the practice of suspending declarations of unconstitutionality.

11.3.4. Roach's Two-Track Remedies Proposal

A promising approach to remedies from which valuable insight can be drawn is one recently propounded by Roach, proposing a 'two-track' model for remedies following judicial review of laws that breach constitutional rights.[25] In his study, Roach draws attention to many of the points canvassed in this thesis, such as how the United States has softened its strong-form review with the as-applied/facial distinction, and how Canada and South Africa have also softened their practice through the use of suspended declarations of invalidity. On foot of this analysis, Roach develops a 'two-track' approach to remedies for unconstitutionality. One track would provide individual litigants with immediate remedies, and the other would seek to address issues more systemically through collaboration with other limbs of government.[26] Different remedies then reflect these different 'tracks': immediate invalidation (as-applied declarations) is litigant-centric, and the declaration of unconstitutionality is systemic.

The advantage of this suggestion, which aims to dispense with the strong- and weak-form distinction in judicial review and take a more nuanced approach,[27] is its flexibility. It allows a court to take a micro- or macro-oriented remedial stance as the case demands. One disadvantage that Roach notes is that this nevertheless leaves open the opportunity for inequity between an individual litigant (who may receive a tailored remedy) and the class of similarly-situated individuals who may receive a substantially different remedy from the legislature, or no remedy at all. Although Roach is sanguine that this difficulty is less than it might appear, his solution amounts to little more than a contingent assertion that we should trust the legislature to do things well in most cases.[28] There are other solutions to this dif-

---

[25] Kent Roach, 'Remedies for Laws that Violate Human Rights' in John Bell and others (eds), *Public Law Adjudication in Common Law Systems* (OUP 2015); Kent Roach, 'Polycentricity and Queue Jumping in Public Law Remedies: A Two-Track Response' (2016) 66 UTLJ 3.

[26] Kent Roach, 'Remedies for Laws that Violate Human Rights' in John Bell and others (eds), *Public Law Adjudication in Common Law Systems* (OUP 2015) 271, 289–96.

[27] ibid 297.

[28] 'The horizontal inequity created by the two-track remedial approach should in most cases be

ficulty, such as the South African interim order, but these raise other difficulties, particularly for the collaborative, systemic track.[29]

It is significant that Roach draws attention to the as-applied challenge and suspended declaration as doctrines that have softened the harder edges of unconstitutionality. These remedies both provide an applicability analysis. Roach suggests that the different remedial tracks will accommodate individual and systemic concerns respectively, but this is difficult to do in the absence of a robust account of applicability. The division between the two tracks is effectively that one type of remedy should provide for the modification of the applicability of an unconstitutional norm, and the other can highlight issues of systemic importance to other institutions. Crudely speaking, this could be thought of as a distinction between applicability (individual remedies) and validity (systemic remedies). This is crude because, as I have argued, applicability is the more significant concept from the point of view of juridical effect; however, it might be that other state institutions will be more alert to being told that a law is 'invalid' than being told that it merely cannot be applied to a discrete range of cases. Most importantly, because applicability can be tailored to a *range* of cases, and not just an individual case, the recognition of this concept may help to address Roach's concern above regarding the inequity between the litigant and similarly-situated parties that can result from remedies on the 'immediate, litigant-focused' track. Courts could draw more fine distinctions in classes of case in which a norm will be unconstitutional; they would not be limited to pronouncing only on the litigant's specific case. In this way, the analytic framework in this thesis can be used to supplement innovative remedial proposals such as that advanced by Roach.

---

temporary. To be sure, the successful litigant will receive a remedy that other similarly situated do not receive, but the whole point of a suspended declaration of invalidity is to give Parliament a year or so to enact new legislation.' Roach, 'Remedies for Laws that Violate Human Rights' (n 26) 296.

[29] I point to some of these in chapter 10. The interim order may defeat the political point of a suspended declaration by nudging the legislature in a particular direction, thus negating the supposed *carte blanche* afforded by the suspended declaration to design a solution to the issue.

11.4. UNCONSTITUTIONALITY AND THE COMPARATIVE METHOD

This section makes some observations applying some findings of this thesis to comparative constitutional law as a discipline more broadly. Leckey has observed that one of many understandings of comparative law is conceptualising it from the perspective of a judge using foreign legal materials.[30] One can also take the perspective of a scholar discussing this judicial perspective.[31] Doyle, in an example of this latter type of scholarship, has argued that declarations of unconstitutionality are among the best candidates for the beneficial application of foreign legal material by judges.[32]

11.4.1. The Use of Foreign Law

Comparative constitutional law has at least two potential audiences: legal practitioners, and academics. Of the legal practitioner class, the audience of judges is of particular importance, as the weight that a judge places on a proposition of foreign law can become authoritative in that judge's home jurisdiction. One of the most significant uses of foreign law is in its ability to influence the outcome of domestic cases. The utility of foreign law in this respect is a function of its reception by the judiciary.

Accounts of precedent will often suggest that foreign law is of 'persuasive' authority. However, as Doyle has observed, what exactly is meant by 'persuasive' in this context is not neatly defined.[33] One might even object that it is a type of category mistake to suggest that any 'authority' can be merely 'persuasive'. To treat a standard as authoritative entails being required to apply or follow it; to treat a stand-

[30] Robert Leckey, 'Review of Comparative Law' (2017) 26 Social & Legal Studies 3, 4.

[31] ibid 4.

[32] Oran Doyle, 'Constitutional Cases, Foreign Law and Theoretical Authority' (2016) 5 Global Constitutionalism 85, 105.

[33] Doyle, 'Constitutional Cases, Foreign Law and Theoretical Authority' (n 32) 86. Citing: HP Glenn, 'Persuasive Authority' (1987) 32 McGill L J 261; Frederick Schauer, 'Authority and Authorities' (2008) 94 Va L Rev 1931.

ard as persuasive means merely accounting for it as a premise in an organic process of reasoning that may reach a different conclusion.[34]

Doyle suggests that this difficulty may be solved by thinking of judges as applying foreign law as *theoretically* authoritative.[35] This is treating something as an authority with a view to knowledge, and can be distinguished from practical authority, which treats a standard as mandatory and normative with a view to action. As Doyle puts it: 'In the same way that practical authority provides an exclusionary reason for action, theoretical authority provides an exclusionary reason for belief.'[36]

As regards unconstitutionality, Doyle claims that this will be one of the least troublesome areas in which judges may turn to foreign law for guidance:

> [T]he most appropriate type of constitutional case for reference to foreign law is one of judicial method. The processes of constitutional litigation produce problems that are generally unanticipated in the constitutional text and best seen as technocratic lawyering issues. The two most obvious of these are questions of standing and the effects of a declaration of unconstitutionality. In this type of case, the problems of misunderstanding are lowest and it is safest to take the foreign case law at face value.[37]

Although Doyle does concede that there can still be differences, and he is correct to characterise unconstitutionality as more technocratic than other areas of constitutional law, the findings of this thesis would suggest that he underestimates the potential for misunderstanding. As the divergences captured in table 8.1 demonstrate, there is considerable variation between jurisdictions, and even within jurisdictions at different times, as to the effects and meaning of unconstitutionality. This can make it anything but safe to take foreign law on unconstitutionality at face value. The confusion that has been generated in the Indian courts through adopting the United States' precedents on unconstitutionality, despite differing constitutional texts and underlying theories, speaks to the perils of embracing foreign precedents too uncritically.

---

[34] For an engagement with this apparent paradox, see: Jeremy Waldron (ed), *Partly Laws Common to All Mankind* (Yale University Press 2012).

[35] Doyle, 'Constitutional Cases, Foreign Law and Theoretical Authority' (n 32) 90.

[36] ibid 90.

[37] ibid 105.

None of this is to say that judges should not have regard to the practice of their foreign brethren.[38] The recent endorsement of suspended declarations of unconstitutionality in Ireland was no doubt inspired by the operation of that remedy in Canada and South Africa. However, there are significant and under-appreciated risks in transplanting practice here. The theoretical framework for unconstitutionality posited in this thesis presents a keystone for the apprehension of similarity and difference that could make the application of foreign law on unconstitutionality more effective in the future.

11.4.2. The Limits of Textual Constraints

One of the salient features of the jurisdictions studied in this thesis is that all of them—bar the United States—ground the practice of judicial review of legislation in the Constitutional text. This fact of itself does not resolve the 'derivation' question, as outlined in chapter 8, definitively. It depends on what the particular Constitution says. A Constitution could direct that it is itself supreme and self-executing (legislative nullity), or it could direct that the judiciary are supreme in deciding judicial review questions (judicial intervention).[39]

A question that is, at least facially, more definitively resolved by the Constitutional texts of these jurisdictions is the 'effects' question. Each text provides for the result to obtain upon a finding of unconstitutionality: in India, such legislation is 'void', in Ireland it is 'invalid', in Canada it is 'of no force or effect'[40] and in South

[38] Indeed, Doyle and Hickey have also more recently drawn attention to the comparatively enlightening application of foreign law by the Irish Supreme Court, including in its unconstitutionality jurisprudence: Oran Doyle and Tom Hickey, 'The Use of Foreign Law in Irish Constitutional Adjudication' in Giuseppe Ferrari (ed), *The Use of Foreign Law by Constitutional Courts* (Brill Publishers 2019).

[39] In both instances the Constitution is 'supreme' in the sense that it provides for the power of judicial review. For the purposes of the derivation question, 'legislative nullity' is meant to identify the constitution's agentive role in constituting invalidity, not the feature of the Constitution that it is the highest-ranking source in the legal system.

[40] Another difficulty here is how similar terms can be interpreted differently in different jurisdictions. The Irish Constitution maintains that pre-constitution laws that are found to be repugnant to the constitution are also 'of no force or effect'. However, the Irish Courts have effectively interpreted this as directing them to regulate the *applicability* of such law. The Canadian courts have

Africa it is 'invalid'. Ignoring developments in case law, on the basis of the analysis in chapter 8, one would expect that the Indian, Irish, and South African systems should endorse a model of invalidity, and the Canadian courts would endorse a model of inapplicability.

However, the limits of the texts as constraints, or even as guides, quickly become apparent in the case law and through comparison. The Canadian courts reject the applicability 'as-applied' analysis and opt instead for an invalidity-based view. This is certainly not required by the text, and, if anything, it makes the practice of suspending declarations more difficult to justify. India is an example that goes the other way: the Indian constitutional text makes it clear that unconstitutional law is void (invalid). However, the Indian courts quickly found that void was a 'relative' term and in *Ambica Mills* it seemed that an applicability-based analysis of unconstitutionality was confirmed. South Africa is perhaps the greatest success story in this regard, as it provides the most comprehensive text on judicial review and the effects of unconstitutionality. However, the point remains that in many cases textual provisions have not provided much meaningful restraint on what unconstitutionality entails.

11.5. CONCLUSION

The framework in this thesis is useful for several reasons. The theoretical questions in chapter 8 can be used to achieve clearer understanding of, and insight to, the practice of unconstitutionality in the jurisdictions under scrutiny. This understanding can supplement existing literature on public law remedies, and even generate theoretical insight into certain specific innovations such as the suspended declaration of unconstitutionality. This can help to reinforce the legitimacy of such innovations, which are becoming increasingly popular.[41] Reorienting the focus of unconstitutionality more generally onto applicability also encourages transparency as regards

interpreted this same phrase as requiring a validity analysis.

[41] See, for example, the growing acceptance of the suspended declaration in Ireland, discussed in chapter 5.

the value-laden judgements that are inevitable in judicial review cases that implicate constitutional rights.

Caution regarding the use of foreign law in unconstitutionality is another important lesson that can be drawn from this work. It has been argued that the effect and praxis of unconstitutionality is highly transferable between jurisdictions. My analysis demonstrates that this is not the case, and that caution should be exercised when regard is had to foreign precedent in developing domestic understanding of unconstitutionality. This is not to say that foreign law is not highly informative and useful, but greater awareness of points of both comparison and contrast is necessary for this exercise to be maximally beneficial.

Finally, the model of unconstitutionality advanced in this thesis as being the most flexible, and the best baseline assumption, is one that conceptualises unconstitutionality as being an exercise of the judicial power, operating on the applicability of legal norms, and presuming prospective effect only. This presumption of prospective effect should, however, be rebuttable and engage in the 'choice of law' style retrospectivity analysis outlined above. This is the model that is most theoretically tenable, and it is also quite remedially flexible. It is a model that builds in a degree of judicial activism that some may find unappealing or concerning. Because the question of retroactivity/prospectivity will be taken on its moral merits in each case, there is a significant degree of discretion and power reserved to the judiciary on this model. However, as this thesis took strong-form judicial review as granted, this weakness is within foreseeable methodological parameters.

In conclusion, therefore, the 'judicial intervention' answer to the derivation question, the 'applicability' answer to the effects question, the 'present' answer to the origin question and the 'retrospective and prospective' answer to the duration question combine to describe the most preferable model of unconstitutionality. Although aspects of this model, particularly its temporal implications, will require modification in some circumstances, the analysis in this thesis would support it as the most appropriate default rule to respond to unconstitutionality as it manifests in the common law tradition.

Bibliography

Articles

Ackerman, Bruce, 'The Rise of World Constitutionalism' (1997) 83 Va L Rev 771

Adler, Matthew, 'Rights Against Rules: The Moral Structure of American Constitutional Law' (1998) 97 Mich L Rev 1

— — 'Rights, Rules, and the Structure of Constitutional Adjudication: A Response to Professor Fallon' (2000) 113 Harv L Rev 1371

Albert, Richard, 'Advisory Review: The Reincarnation of the Notwithstanding Clause' [2008] Alberta L Rev 1037

— — 'Constitutional Amendment by Constitutional Desuetude' (2014) 62 Am J Comp Law 641

— — 'Constitutional Disuse or Desuetude: The Case of Article V' (2014) 94 Boston Uni L Rev 1029

Beytagh, Francis, 'Ten Years of Non-Retroactivity: A Critique and a Proposal' (1975) 61 Va L Rev 1557

Bix, Brian, 'Global Error and Legal Truth' (2009) 29 OJLS 535

Black, Oliver, 'Legal Validity and the Infinite Regress' (1996) 15 Law and Philosophy 339

Buck, Stuart and Rienzi, Mark, 'Federal Courts, Overbreadth, and Vagueness: Guiding Principles for Constitutional Challenges to Uninterpreted State Statutes' [2002] Utah L Rev 381

Bulygin, Eugenio, 'An Antinomy in Kelsen's Pure Theory of Law' (1990) 3 Ratio Juris 29

Bulygin, Eugenio, 'Algunas Consideraciones Sobre los Sistemas Jurídicos' (1991) 9 Doxa 257

Burazin, Luka, 'Can There Be an Artifact Theory of Law?' (2016) 29 Ratio Juris 385

Cappelletti, Mauro, 'Judicial Review in Comparative Perspective' (1970) 58 Ca L Rev 1017

Carolan, Eoin, 'The relationship between judicial remedies and the separation of powers: collaborative constitutionalism and the suspended declaration of invalidity' (2011) 46 Ir Jur 180

— — 'Dialogue Isn't Working: The Case for Collaboration as Model of Legislative–Judicial Relations' (2016) 36 Legal Studies 209

Chandrachud, Chintan, 'Declarations of Unconstitutionality in India and the UK: Comparing the Space for Political Response' (2015) 43 Ga J Int'l & Comp L 309

Choudhry, Sujit and Roach, Kent, 'Putting the Past Behind Us? Prospective Judicial and Legislative Constitutional Remedies' (2003) 21 Supreme Court Law Review (2nd) 205

Clark, Michael and Liggins, David, 'Recent Work on Grounding' 72 Analysis 812

Clarke, Desmond, 'Judicial Reasoning: Logic, Authority, and the Rule of Law in Irish Courts' (2011) 46 Ir Jur 152

Claus, Laurence, 'Montesquieu's Mistakes and the True Meaning of Separation' (2005) 25 OJLS 419

Coleman, Jules, 'Truth and Objectivity in Law' (1995) 1 Legal Theory 33

— — 'The Architecture of Jurisprudence' (2011) 121 Yale L J 2

Coleman, Jules and Leiter, Brian, 'Determinacy, Objectivity, and Authority' (1993) 142 U Pa L Rev 549

Comella, Victor Ferreres, 'The European model of constitutional review of legislation: Toward decentralization?' (2004) 2 International Journal of Constitutional Law 461

Corr, John, 'Retroactivity: A Study in Supreme Court Doctrine as Applied' (1983) 61 N C L Rev 745

Coyle, Seán, 'Hart, Raz, and the Concept of a Legal System' (2002) 21 Law & Philosophy 275

Crawford, Earl, 'The Legislative Status of an Unconstitutional Statute' (1951) 49 Mich L Rev 645

Crowe, Jonathan, 'Law as an Artifact Kind' (2014) 40 Monash U L Rev 737

Currier, Thomas, 'Time and Change in Judge-Made Law: Prospective Overruling' (1965) 51 Va L Rev 201

de Andrade, Gustavo Fernandes, 'Comparative Constitutional Law: Judicial Review' (2001) 3 U Pa J Const L 977

De Moor, Anne, 'Nothing Else to Think? On Meaning, Truth, and Objectivity in Law' (1998) 18 OJLS 345

Dewey, John, 'Logical Method and Law' (1924) 10 Cornell L Rev 17

Dorf, Michael, 'Facial Challenges to State and Federal Statutes' (1994) 46 Stan L Rev 235

Doyle, Oran, 'Conventional Constitutional Law' (2015) 38 DULJ 311

— — 'Constitutional Cases, Foreign Law and Theoretical Authority' (2016) 5 Global Constitutionalism 85

Dummett, Michael, 'Bringing About the Past' (1964) 73 The Philosophical Review 338

Dworkin, Ronald, 'Hard Cases' (1975) 88 Harv L Rev 1057

Easterbrook, Frank, 'Presidential Review' (1990) 40 Case W Res L Rev 905

Fallon, Richard, 'Making Sense of Overbreadth' (1991) 100 Yale L J 876

— — 'As-Applied and Facial Challenges and Third-Party Standing' (2000) 113 Harv L Rev 1321

— — 'Fact and Fiction about Facial Challenges' (2011) 99 Cal L Rev 914

Fallon, Richard and Meltzer, Daniel, 'New Law, Non-Retroactivity, and Constitutional Remedies' (1991) 104 Harv L Rev 1731

Fanning, Rossa, 'Hard Case; Bad Law? The Supreme Court Decision in *A v Governor of Arbour Hill Prison*' (2005) 40 Ir Jur 188

Feldman, David, 'Error of Law and Flawed Administrative Acts' (2014) 73 CLJ 275

Finnis, John, 'On Reason and Authority in "Law's Empire"' (1987) 6 Law and Philosophy 357

— — 'The Fairy Tale's Moral' (1999) 115 LQR 170

Fish, Eric, 'Severability as Conditionality' (2015) 64 Emory L J 1293

— — 'Choosing Constitutional Remedies' (2016) 63 UCLA L Rev 322

Fitzgerald, Brian, 'When Should Unconstitutionality Mean "Void Ab Initio"' (1994) 1 Canberra L Rev 205

Foley, Kathleen, 'Australian Judicial Review' (2007) 6 Wash U Global Stud L Rev 281

Frankenberg, Günter, 'Critical Comparisons: Re-thinking Comparative Law' (1985) 26 Harv Int'l L J 411

— — 'The Innocence of Method – Unveiled: Comparison as an Ethical and Political Act' (2014) 9 J Comp L 222

Franklin, David, 'Looking Through Both Ends of the Telescope: Facial Challenges and the Roberts Court' (2009) 36 Hastings Const L Q 689

Fuller, Lon, 'The Forms and Limits of Adjudication' (1978) 92 Harv L Rev 353

Geller, Joshua, 'Truth, Objectivity, and Dworkin's Right Answer Thesis' [1999] UCL Jurisprudence Rev 83

Gkouvas, Triantafyllos, 'Resisting Perspectivalism About Law: The Scope of Jurisprudential Disagreement' (2017) 8 Jurisprudence 205

Glenn, HP, 'Persuasive Authority' (1987) 32 McGill L J 261

Gorman, Robert, 'Supreme Court Judgment of State Statute as Unconstitutional on its Face' (1956) 31 Notre Dame L Rev 684

Graetz, Michael, 'Retroactivity Revisited' (1985) 98 Harv L Rev 1820

Greenberg, Mark, 'The Moral Impact Theory of Law' (2014) 123 Yale L J 1289

Grellette, Matthew, 'Legal Positivism and the Separation of Existence and Validity' (2010) 23 Ratio Juris 22

Guidice, Michael, 'Unconstitutionality, Invalidity, and Charter Challenges' (2002) 15 Can J L & Jurisprudence 69

Haddad, James, '"Retroactivity Should be Rethought": A Call for the End of the *Linkletter* Doctrine' (1969) 60 The Journal of Criminal Law, Criminology, and Police Science 417

Hammer, Stephen, 'Retroactivity and Restraint: An Anglo-American Comparison' (2018) 41 Harv J L & Pub Pol'y 409

Harding, Maebh, 'The Curious Incident of the Marriage Act (No 2) 1537 and the Irish Statute Book' (2012) 32 Legal Stud 78

Hartnett, Edward, 'Modest Hope for a Modest Roberts Court: Deference, Facial Challenges, and the Comparative Competence of Courts' (2006) 59 SMU L Rev 1735

Hasler, Pierce, 'Retroactivity Rethought: The Hidden Costs' (1972) 24 Me L Rev 1

Henriques, Mark, 'Desuetude and Declaratory Judgment: A New Challenge to Obsolete Laws' (1990) 76 Va L Rev 1057

Higgins, Rosalyn, 'Time and the Law: International Perspectives on an Old Problem' (1997) 46 ICLQ 501

Himma, Kenneth, 'Final Authority to Bind With Moral Mistakes: On the Explanatory Potential of Inclusive Legal Positivism' (2005) 24 Law & Philosophy 1

Hogan, Gerard, 'Statutory interpretation – the doctrine of desuetude' (1987) 9 DULJ 136

Hogg, Peter, 'Necessity in a Constitutional Crisis' (1985) 15 Monash U L Rev 253

Hogg, Peter, Bushell Thornton, Allison and Wright, Wade, 'Charter Dialogue Revisited: Or "Much Ado About Metaphors"' (2007) 45 Osgoode Hall L J 1

Hogg, Peter and Bushell, Allison, 'The Charter Dialogue between Courts and Legislatures (Or Perhaps the Charter of Rights Isn't Such a Bad Thing after All)' (1997) 35 Osgoode Hall L J 75

Hogg, Peter and Zwibel, Cara, 'The Rule of Law in the Supreme Court of Canada' (2005) 55 UTLJ 715

Houkes, Wybo and Vermaas, Pieter, 'Actions Versus Functions: A Plea for an Alternative Metaphysics of Artifacts' (2004) 87 The Monist 52

Isserles, Mark, 'Overcoming Overbreadth: Facial Challenges and the Valid Rule Requirement' (1998) 48 Am U L Rev 359

Kaplow, Louis, 'An Economic Analysis of Legal Transactions' (1986) 99 Harv L Rev 509

Katyal, Neal and Schmidt, Thomas, 'Active Avoidance: The Modern Supreme Court and Legal Change' (2015) 128 Harv L Rev 2109

Kavanagh, Aileen, 'The Idea of a Living Constitution' (2003) 16 Can J L & Jurisprudence 55

— — 'A Hard Look at the Last Word' (2015) 35 OJLS 825

— — 'The Lure and the Limits of Dialogue' (2016) 66 UTLJ 83

Kay, Richard, 'Retroactivity and Prospectivity of Judgments in American Law' (2014) 62 American Journal of Comparative Law 37

Keller, Scott and Tesytlin, Misha, 'Applying Constitutional Decision Rules Versus Invalidating Statutes *in Toto*' (2012) 98 Va L Rev 301

Kelsen, Hans, 'Judicial Review of Legislation' (1942) 4 Journal of Politics 183

Kenny, David, 'Grounding Constitutional Remedies in Reality: The Case for as-Applied Constitutional Challenges in Ireland' (2014) 37 DULJ 53

Klukowski, Kenneth, 'Severability Doctrine: How Much of a Statute Should Federal Courts Invalidate' (2011) 16 Tex Rev L & Pol 1

Koviacic, Candice, 'Remedying Underinclusive Statutes' (1986) 33 Wayne L Rev 39

Kramer, Matthew, 'Is Law's Conventionality Consistent with Law's Objectivity?' (2008) 14 Res Publica 241

Kress, Kenneth, 'Legal Reasoning and Coherence Theories: Dworkin's Rights Thesis, Retroactivity, and the Linear Order of Decisions' (1984) 72 CLR 369

Law, David, 'The Anatomy of a Conservative Court: Judicial Review in Japan' (2008) 87 Tex L Rev 1545

— — 'Why has Judicial Review Failed in Japan' (2010) 88 Wash U L Rev 1425

Lawlor, Ciarán, 'Troubling Times: Intertemporal Law and Theories of Approach to the Effects of Unconstitutionality' (2008) 7 COLR 71

Leckey, Robert, 'Enforcing Laws That Infringe Rights' [2016] PL 206

— — 'Remedial Practice Beyond Constitutional Text' (2016) 64 Am J Comp L 1

— — 'The Harms of Remedial Discretion' (2016) 14 ICON 584

— — 'Review of Comparative Law' (2017) 26 Social & Legal Studies 3

Leiter, Brian, 'The Demarcation Problem in Jurisprudence: A New Case for Skepticism' (2011) 31 OJLS 663

MacCormack, G, '"Law" and "Legal System"' (1979) 42 MLR 285

McCreary, Jackie, 'Retroactivity of Laws: An Illustration of Intertemporal Conflicts Law Issues through the Revised Civil Code Articles on Disinherison' (2002) 62 La L Rev 1321

McMahon, Bryan, 'Review of JM Kelly: The Irish Constitution' (1981) 16 Ir Jur 190

McManus, Richard, 'Predicting the Past: The Declaratory Theory of the Common Law –From Fairytale to Nightmare' (2007) 12 Judicial Review 228

McNulty, John, 'Corporations and the Intertemporal Conflict of Laws' (1967) 55 Ca L Rev 12

Meier, Luke, 'Facial Challenges and Separation of Powers' (2010) 85 Ind L J 1557

Metzger, Gillian, 'Facial Challenges and Federalism' (2005) 105 Colum L Rev 873

Mezey, Susan Gluck, 'Civil Law and Common Law Traditions: Judicial Review and Legislative Supremacy in West Germany and Canada' [1983] International and Comparative Law Quarterly 689

Miller, Bradley, 'Beguiled by Metaphors: The Living Tree and Originalist Constitutional Interpretation in Canada' (2009) 22 Can J L & Jurisprudence 331

Mirfield, Peter, 'A Challenge to the Declaratory Theory of Law' (2008) 124 LQR 190

Mishkin, Paul, 'Foreword: The High Court, the Great Writ, and the Due Process of Time and Law' (1965) 79 Harv L Rev 56

Mitchell, Jonathan, 'The Writ-of-Erasure Fallacy' (2018) 104 Va L Rev 933

Monaghan, Henry, 'Overbreadth' [1981] S Ct Rev 1

Moore, Michael, 'The Semantics of Judging' (1981) 54 S Cal L Rev 151

— — 'A Natural Law Theory of Interpretation' (1985) 58 Cal L Rev 277

— — 'Metaphysics, Epistemology and Legal Theory' (1987) 60 S Cal L Rev 453

Moore, Michael, 'The Interpretive Turn in Modern Theory: A Turn for the Worse?' (1989) 41 Stan L Rev 871

— — 'Justifying the Natural Law Theory of Constitutional Interpretation' (2001) 69 Fordham L Rev 2087

— — 'Legal Reality: A Naturalist Approach to Legal Ontology' (2002) 21 Law & Philosophy 619

Mouland, Carolyn, 'Remedying the Remedy: *Bedford*'s Suspended Declaration of Invalidity' (2018) 41 Manitoba L J 281

Munzer, Stephen, 'Validity and Legal Conflicts' (1973) 82 Yale L J 1140

— — 'Retroactive Law' (1977) 6 J Legal Stud 373

— — 'Realistic Limits on Realist Interpretation' (1985) 58 S Cal L Rev 459

Murphy, William Mark, 'The Problem of Unconstitutionality and Retroactivity in Criminal Law: Ireland, the US and Canada Compared' (2007) 42 Ir Jur 63

Navarro, Pablo and Moreso, José Juan, 'Applicability and Effectiveness of Legal Norms' (1997) 16 Law & Philosophy 201

Navarro, Pablo and others, 'Applicability of Legal Norms' (2004) 17 Can J L & Jurisprudence 337

Nimmer, Melville, 'A Proposal for Judicial Validation of a Previously Unconstitutional Law: The Civil Rights Act of 1875' (1965) 65 Colum L Rev 1394

Noonan, Robert, '*Stare Decisis*, Overruling, and Judicial Law-Making: The Paradox of the *JC* Case' (2017) 57 Ir Jur 119

O'Sullivan, Julie, '*United States v Johnson*: Reformulating the Retroactivity Doctrine' (1983) 69 Cornell L Rev 166

Patterson, Dennis, 'Realist Semantics and Legal Theory' (1989) 2 Can J L & Jurisprudence 175

— — 'Semantics and Legal Interpretation (Further Thoughts)' (1989) 2 Can J L & Jurisprudence 181

Pino, Giorgio, 'Farewell to the Rule of Recognition?' (2011) 5 Problema 265

Raz, Joseph, 'The Identity of Legal Systems' (1971) 59 Ca L Rev 795

Reid, Lord, 'The Judge as Lawmaker' (1972–1973) 12 J Soc Public Teachers 22

Rheinstein, Max, 'Place of Wrong: A Study in the Method of Case Law' (1944) 19 Tulane L Rev 4

Ring, Sinéad, 'Collateral Challenges to a Declaration of Unconstitutionality: *A v Governor of Arbour Hill Prison*' (2006) 14 ISLR 245

Roach, Kent, 'Enforcement of the Charter—Subsections 24(1) and 52(1)' (2013) 62 SCLR (2d) 473

— — 'Polycentricity and Queue Jumping in Public Law Remedies: A Two-Track Response' (2016) 66 UTLJ 3

Roache, Rebecca, 'What is it Like to Affect the Past?' (2015) 34 Topoi 195

Rodgers, Linda and Rodgers, William, 'Desuetude as a Defence' (1966) 52 Iowa L Rev 1

Roosevelt, Kermit, 'A Little Theory is a Dangerous Thing: The Myth of Adjudicative Retroactivity' (1999) 31 Conn L Rev 1075

Rosati, Connie, 'Some Puzzles About the Objectivity of Law' (2004) 23 Law & Philosophy 273

Rossum, Ralph, 'New Rights and Old Wrongs: The Supreme Court and the Problem of Retroactivity' (1974) 23 Emory L J 381

Ruiter, Dick, 'Legal Validity Qua Specific Mode of Existence' (1997) 16 Law & Philosophy 579

Ryder, Bruce, 'Suspending the Charter' (2003) 21 Supreme Court Law Review 267

Schafer-Landau, Russ, 'Moral Rules' (1997) 107 Ethics 584

Schauer, Frederick, 'Authority and Authorities' (2008) 94 Va L Rev 1931

Schrödinger, Erwin, 'Die gegenwärtige Situation in der Quantenmechanik' (1935) 23(48) Naturwissenschaften 807

Schwartz, Stephen, 'Putnam on Artifacts' (1978) 87 Philosophical Review 566

Shapiro, David, 'State Courts and Federal Declaratory Judgments' (1979) 74 Nw U L Rev 759

Sider, Theodore, 'Presentism and Ontological Commitment' (1999) 96 Journal of Philosophy 325

Skeffington, Patricia Sheehy, 'Commencement Orders: A Creeping Incursion into the Legislature's Law-Making Domain?' (2018) 59 Ir Jur 93

Smith, Jennifer, 'The Origins of Judicial Review in Canada' (1983) 16 Canadian Journal of Political Science 115

Solum, Lawrence, 'The Fixation Thesis: The Role of Historical Fact in Original Meaning' (2015) 91 Notre Dame L Rev 1

Swaminathan, Shivprasad, 'Schrödinger's Constitutional Cat: Limits of the High Court's Declaration of Unconstitutionality' (2013) 25 Nat'l L Sch India Rev 100

Torcia, Charles and King, Donald, 'The Mirage of Retroactivity and Changing Constitutional Concepts' (1962) 66 Dick L Rev 269

Trahan, John Randall, 'Time for a Change: A Call to Reform Louisiana's Intertemporal Conflicts Law (Law of Retroactivity of Laws)' (1999) 59 La L Rev 661

Treanor, William and Sperling, Gene, 'Prospective Overruling and the Revival of "Unconstitutional" Statutes' (1993) 93 Colum L Rev 1902

Tushnet, Mark, 'Policy Distortion and Democratic Debilitation: Comparative Illumination of the Countermajoritarian Difficulty' (1995) 94 Mich L Rev 245

Books

Alchourrón, Carlos and Bulygin, Eugenio, *Normative Systems* (Springer-Verlag 1971)

Austin, John, *Lectures on Jurisprudence* (vol 2, John Murray 1863)

Basu, Durga Das, *Commentary on the Constitution of India* (8th edn, LexisNexis 2011)

Bernal, Carlos and others (eds), *Essays in Legal Philosophy* (OUP 2015)

Bix, Brian, *Law, Language and Legal Determinacy* (Clarendon 1995)

— — *Jurisprudence: Theory and Context* (6th edn, Sweet & Maxwell 2012)

Blackstone, William, *Commentaries on the Laws of England* (vol 1, Cavendish 1766)

Briggs, Adrian and others (eds), *Dicey, Morris and Collins on The Conflict of Laws* (15th edn, Thomson Reuters 2012)

Cardozo, Benjamin, *The Nature of the Judicial Process* (Yale University Press 1921)

Casey, James, *Constitutional Law in Ireland* (3rd edn, Sweet & Maxwell 2000)

Cooley, Thomas, *A Treatise on the Constitutional Limitations which Rest Upon the Legislative Power of the States of the American Union* (7th edn, first published 1868, Little, Brown & Co 1903)

Craig, Paul, *Administrative Law* (7th edn, Sweet & Maxwell 2012)

Cruz, Isagani, *Constitutional Law* (Central Law Book Publishing 1991)

Currie, Iain and de Wall, Johan (eds), *The Bill of Rights Handbook* (5th edn, Juta 2005)

Delaney, Erin and Dixon, Rosalind (eds), *Comparative Judicial Review* (Edward Elgar 2018)

Doyle, Oran, *Constitutional Law: Text, Cases and Materials* (Clarus Press 2008)

Dworkin, Ronald, *Taking Rights Seriously* (Duckworth 1977)

— — *A Matter of Principle* (Harvard University Press 1985)

— — *Law's Empire* (Hart Publishing 1998)

Field, Oliver P, *The Effect of an Unconstitutional Statute* (University of Minnesota Press 1935)

Gardner, John, *Law as a Leap of Faith* (OUP 2012)

Ginsburg, Tom and Dixon, Rosalind (eds), *Comparative Constitutional Law* (Edward Elgar 2013)

Hart, HLA, *The Concept of Law* (2nd edn, OUP 1994)

Hidayatullah, Mohammed, *Constitutional Law of India* (vol 1, Lucas Publications 1986)

Hogan, Gerard and Whyte, Gerry, *J M Kelly: The Irish Constitution* (4th edn, Tottel Publishing 2003)

Hogg, Peter, *Constitutional Law of Canada* (5th edn, Carswell 2007)

Jackson, Vicki and Tushnet, Mark, *Comparative Constitutional Law* (2nd edn, Foundation Press 2006)

Kelsen, Hans, *General Theory of Law and State* (Wedberg, Anders tr, Harvard University Press 1945)

— — *Pure Theory of Law* (Knight, Max tr, University of California Press 1967)

Leckey, Robert, *Bills of Rights in the Common Law* (CUP 2015)

Leiter, Brian, *Naturalizing Jurisprudence: Essays on American Legal Realism and Naturalism in Legal Philosophy* (OUP 2007)

MacCormick, Neil, *Institutions of Law: An Essay in Legal Theory* (OUP 2007)

Marmor, Andrei, *Interpretation and Legal Theory* (OUP 1992)

Munzer, Stephen, *Legal Validity* (Martinus Njihoff 1972)

Navarro, Pablo and Rodríguez, Jorge, *Deontic Logic and Legal Systems* (Cambridge University Press 2014)

Popelier, Patricia, Verstraelen, Sarah and Vanheule, Dirk (eds), *The Effects of Judicial Decisions in Time* (Intersentia 2013)

Rautenbach, IM and Malherbe, EFJ, *Constitutional Law* (LexisNexis Butterworths 2004)

Raz, Joseph, *The Concept of a Legal System: An Introduction to the Theory of Legal System* (Clarendon Press 1980)

— — *Ethics in the Public Domain* (rev edn, OUP 1995)

— — *The Authority of Law: Essays on Law and Morality* (2nd edn, OUP 2009)

Rosenfeld, Michael and Sajo, Andras (eds), *The Oxford Handbook of Comparative Constitutional Law* (OUP 2012)

Rudder Baker, Lynne, *The Metaphysics of Everyday Life: An Essay in Practical Realism* (Cambridge University Press 2007)

Sampford, Charles, *Retrospectivity and the Rule of Law* (OUP 2006)

Stavropoulos, Nicos, *Objectivity in Law* (OUP 1996)

Steiner, Eva (ed), *Comparing the Prospective Effect of Judicial Rulings Across Jurisdictions* (Springer 2015)

Tushnet, Mark, *Weak Courts, Strong Rights: Judicial Review and Social Welfare Rights in Comparative Constitutional Law* (Princeton University Press 2008)

— — *Advanced Introduction to Comparative Constitutional Law* (Edward Elgar 2014)

von Wright, Georg Henrik, *Norm and Action: A Logical Enquiry* (Routledge & Kegan Paul 1963)

Wade, Henry, *Administrative Law* (3rd edn, Clarendon Press 1971)

Waldron, Jeremy (ed), *Partly Laws Common to All Mankind* (Yale University Press 2012)

Waluchow, Wil, *Inclusive Legal Positivism* (Clarendon Press 1994)

Willoughby, Westel Woodbury, *The Constitutional Law of the United States* (Baker, Voorhis & Co 1910)

Woolman, Stuart and Bishop, Michael (eds), *Constitutional Law of South Africa* (2nd edn (rev), Juta 2013)

Contributions to Collections

Albert, Richard, 'The Desuetude of the Notwithstanding Clause – And How to Revive It' in Macfarlane, Emmett (ed), *Policy Change, Courts, and the Canadian Constitution* (University of Toronto Press 2018)

Bix, Brian, 'Metaphysical Realism and Legal Reasoning: The Philosophy of Michael S Moore' in Ferzan, Kimberly and Morse, Stephen (eds), *Legal, Moral, and Metaphysical Truths* (OUP 2016)

Bulygin, Eugenio, 'Judicial Decisions and the Creation of Law' in Bernal, Carlos and others (eds), *Essays in Legal Philosophy* (OUP 2015)

— — 'Kelsen on the Completeness and Consistency of Law', in Bernal, Carlos and others (eds), *Essays in Legal Philosophy* (OUP 2015)

— — 'Norms, Normative Propositions, and Legal Statements', in Bernal, Carlos and others (eds), *Essays in Legal Philosophy* (OUP 2015)

— — 'The Expressive Conception of Norms', in Bernal, Carlos and others (eds), *Essays in Legal Philosophy* (OUP 2015)

— — 'The Problem of Legal Validity in Kelsen's Pure Theory of Law', in Bernal, Carlos and others (eds), *Essays in Legal Philosophy* (OUP 2015)

— — 'Time and Validity', in Bernal, Carlos and others (eds), *Essays in Legal Philosophy* (OUP 2015)

Doyle, Oran and Hickey, Tom, 'The Use of Foreign Law in Irish Constitutional Adjudication' in Ferrari, Giuseppe (ed), *The Use of Foreign Law by Constitutional Courts* (Brill Publishers 2019)

Dworkin, Ronald, 'No Right Answer?' in Hacker, Peter and Raz, Joseph (eds), *Law, Morality and Society: Essays in Honour of HLA Hart* (Clarendon Press 1977)

— — 'On Gaps in the Law', in MacCormick, Neil and Amselek, Paul (eds), *Controversies About Law's Ontology* (Edinburgh University Press 1991)

Endicott, Timothy, 'Raz on Gaps—The Surprising Part' in Meyer, Lukas, Paulson, Stanley and Pogge, Thomas (eds), *Rights, Culture and the Law: Themes from the Legal and Political Philosophy of Joseph Raz* (OUP 2003)

Finnis, John, 'Revolutions and Continuity of Law' in Simpson, AWB (ed), *Oxford Essays in Jurisprudence: Second Series* (OUP 1972)

Forsyth, Christopher, '"The Metaphysic of Nullity": Invalidity, Conceptual Reasoning and the Rule of Law' in Forsyth, Christopher and Hare, Ivan (eds), *The Golden Metwand and the Crooked Cord: Essays in Honour of Sir William Wade QC* (OUP 1998)

Gardner, John, 'Can There be a Written Constitution?' in Green, Leslie and Leiter, Brian (eds), *Oxford Studies in Philosophy of Law: Volume 1* (OUP 2011)

— — 'Legal Positivism: $5\frac{1}{2}$ Myths', in Gardner, John, *Law as a Leap of Faith* (OUP 2012)

— — 'The Legality of Law', in Gardner, John, *Law as a Leap of Faith* (OUP 2012)

Green, Leslie, 'The Morality in Law' in Duarte d'Almeida, Luis, Edwards, James and Dolcetti, Andrea (eds), *Reading HLA Hart's* The Concept of Law (OUP 2013)

Greenberg, Mark, 'How Facts Make Law I' in Hershovitz, Scott (ed), *Exploring Law's Empire: The Jurisprudence of Ronald Dworkin* (OUP 2008)

Hart, HLA, 'Positivism and the Separation of Law and Morals' in Hart, HLA, *Essays in Jurisprudence and Philosophy* (Clarendon Press 1983)

Hogan, Gerard, 'Declarations of Incompatibility, Inapplicability and Invalidity: Rights, Remedies and the Aftermath' in Bradley, Kieran, Travers, Noel and Whelan, An-

thony (eds), *Of Courts and Constitutions:* Liber Amicorum *in Honour of Niall Fennelly* (Hart Publishing 2014)

Jackson, Vicki, 'Comparative Constitutional Law: Methodologies' in Rosenfeld, Michael and Sajo, Andras (eds), *The Oxford Handbook of Comparative Constitutional Law* (OUP 2012)

Kenny, David, 'The Separation of Powers and Remedies: The Legislative Power and Remedies for Unconstitutional Legislation in Comparative Perspective' in Carolan, Eoin (ed), *The Constitution of Ireland: Perspectives and Prospects* (Bloomsbury Professional 2012)

Leiter, Brian, 'Law and Objectivity' in Leiter, Brian, *Naturalizing Jurisprudence: Essays on American Legal Realism and Naturalism in Legal Philosophy* (OUP 2007)

— — 'Objectivity, Morality, and Adjudication', in Leiter, Brian, *Naturalizing Jurisprudence: Essays on American Legal Realism and Naturalism in Legal Philosophy* (OUP 2007)

— — 'Legal Positivism about the Artifact Law: A Retrospective Assessment', in Burazin, Luka, Himma, Kenneth and Roversi, Corrado (eds), *Law as an Artifact* (OUP 2017)

Moore, Michael, 'Law as a Functional Kind' in George, Robert (ed), *Natural Law Theory: Contemporary Essays* (OUP 1992)

— — 'Truth-makers for Propositions of Law and for Propositions About Law', in Keil, Geert and Poscher, Ralf (eds), *Vagueness and the Law* (OUP 2014)

Niblett, Anthony, 'Delaying Declarations of Constitutional Invalidity' in Fagan, Frank and Levmore, Saul (eds), *The Timing of Lawmaking* (Edward Elgar 2017)

Penner, James, 'Law and Adjudication: Dworkin's Critique of Positivism' in Penner, James, Schiff, David and Nobles, Richard (eds), *Jurisprudence & Legal Theory: Commentary and Materials* (OUP 2002)

Raz, Joseph, 'On the Authority and Interpretation of Constitutions: Some Preliminaries' in Alexander, Larry (ed), *Constitutionalism: Philosophical Foundations* (Cambridge University Press 1998)

Raz, Joseph, 'Law and Value in Adjudication' in Raz, Joseph, *The Authority of Law: Essays on Law and Morality* (2nd edn, OUP 2009)

— — 'Legal Reasons, Sources, and Gaps', in Raz, Joseph, *The Authority of Law: Essays on Law and Morality* (2nd edn, OUP 2009)

— — 'Legal Validity', in Raz, Joseph, *The Authority of Law: Essays on Law and Morality* (2nd edn, OUP 2009)

— — 'Postema on Law's Autonomy and Public Practical Reasons: A Critical Comment', in Raz, Joseph, *Between Authority and Interpretation* (OUP 2009)

— — 'The Institutional Nature of Law', in Raz, Joseph, *The Authority of Law: Essays on Law and Morality* (2nd edn, OUP 2009)

Roach, Kent, 'Remedies for Laws that Violate Human Rights' in Bell, John and others (eds), *Public Law Adjudication in Common Law Systems* (OUP 2015)

Rosen, Gideon, 'Metaphysical Dependence: Grounding and Reduction' in Hale, Bob and Hoffmann, Aviv (eds), *Modality: Metaphysics, Logic, and Epistemology* (OUP 2010)

Soames, Scott, 'What Vagueness and Indeterminacy Tell Us About Interpretation' in Marmor, Andrei and Soames, Scott (eds), *The Philosophical Foundations of Language in the Law* (OUP 2011)

Trogdon, Kelly, 'An Introduction to Grounding' in Hoeltje, Miguel, Schnieder, Benjamin and Steinberg, Alex (eds), *Varieties of Dependence: Ontological Dependence, Grounding, Supervenience, Response-Dependence (Basic Philosophical Concepts)* (Philosophia Verlag 2013)

Tushnet, Mark, 'The Rise of Weak-Form Judicial Review' in Ginsburg, Tom and Dixon, Rosalind (eds), *Comparative Constitutional Law* (Edward Elgar 2011)

Waluchow, Wil, 'Four Concepts of Legal Validity' in Adler, Matthew and Himma, Kenneth (eds), *The Rule of Recognition and the US Constitution* (OUP 2009)

Woozley, Anthony, 'No Right Answer' in Cohen, Marshall (ed), *Ronald Dworkin and Contemporary Jurisprudence* (Duckworth 1984)

Other Works

American Jurisprudence (5th edn)

Dorf, Michael, 'Whither Severability After *Murphy v NCAA*' (*Dorf on Law*, 17th May 2018) ⟨http://www.dorfonlaw.org/2018/05/whither-severability-after-murphy-v-ncaa.html⟩ accessed 11th June 2018

Green, Leslie 'Legal Positivism', *The Stanford Encyclopedia of Philosophy* (Fall edn, 2009) ⟨http://plato.stanford.edu/archives/fall2009/entries/legal-positivism/⟩ accessed 5th December 2014

Hoole, Grant, 'Proportionality as a Remedial Principle: A Framework for Suspended Declarations of Invalidity in Canadian Constitutional Law' (LLM thesis, 2010)

Leckey, Robert, 'Suspended Declarations of Invalidity and the Rule of Law' (*UK Const Law Blog*, 14th March 2014) ⟨https://ukconstitutionallaw.org/2014/03/12/robert-leckey-suspended-declarations-of-invalidity-and-and-the-rule-of-law/⟩ accessed 21st March 2017

Marín, Rafaél Hernández and Sartor, Giovanni, 'Time and Norms: A Formalisataion in the Event-Calculus' (ICAIL '99 Proceedings of the 7th International Conference on Artificial intelligence and Law, Oslo, June 1999)

Rhodes, Charles, 'Over the Threshold of Constitutional Adjudicative Retroactivity' (*SSRN Draft Paper*, 24th May 2016) ⟨https://ssrn.com/abstract=2783954⟩ accessed 18th December 2017

Weisstein, Eric 'Totally Ordered Set', *MathWorld—A Wolfram Web Resource* ⟨http://mathworld.wolfram.com/TotallyOrderedSet.html⟩ accessed 12th October 2016

Wenar, Leif 'Rights', *The Stanford Encyclopedia of Philosophy* (Fall edn, 2015) ⟨https://plato.stanford.edu/archives/fall2015/entries/rights/⟩ accessed 24th March 2017

Wong, Julia and Heidenreich, Phil, 'Federal justice minister moves to end "zombie laws," clean up Criminal Code' (*Global News*, 8th March 2017) ⟨https://

globalnews . ca / news / 3295578 / federal - government - expected - to - remove - zombie-laws-from-criminal-code/⟩ accessed 3rd December 2017