

MANCHESTER 1824

The University of Manchester

REVIEW OF LAW, CRIME AND ETHICS

VOLUME III | OCTOBER 2014

The logo for Jones Day, featuring the words "JONES" and "DAY" stacked vertically in a white, serif font. A registered trademark symbol (®) is located at the bottom right of the word "DAY". The logo is set against a dark blue square background.

JONES
DAY®

The Manchester Review of Law, Crime & Ethics

The University of Manchester's
School of Law Student Journal

VOL. III

The Manchester Review of Law, Crime & Ethics is a student-led, peer-reviewed journal founded at the University of Manchester, School of Law. The journal exhibits the best academic work in law, criminology and ethics, on both the undergraduate and postgraduate levels, publishing it annually.

VOLUME III

OCTOBER 2014

The Manchester Review of Law, Crime & Ethics

© *The Manchester Review of Law, Crime & Ethics*, 2014
All rights reserved

The *Manchester Review of Law, Crime & Ethics* and its editors are not responsible for the opinions expressed in this journal.

No part of this journal may be reproduced in any form without the express consent of The *Manchester Review of Law, Crime & Ethics* or the University of Manchester, School of Law.

All enquiries should be addressed to:

Manchester Review of Law, Crime & Ethics
Williamson Building, School of Law
University of Manchester
Manchester
M13 9PL

review.manchester@gmail.com

Typeset in Baskerville Old Face
Printed in United Kingdom

The Manchester Review of Law, Crime & Ethics

Editor-in-Chief

Joe Tomlinson

Deputy Editor-in-Chief

Emile Abdul-Wahab

Editors

Amin Al-Atsewani

Philip Burton

Dorota Galeza

Janus Lim

Matthew Lowry

Amelia McGrath

Daniel Watson

Hannah Wishart

Max Melsa

Editorial Assistance

Sonam Cheema

Koryann Stevens-Delves

Elizabeth O'Loughlin

Honorary Advisory Board

Lesley Anderson QC

Dr Nicholas Braslavsky

HH Judge Nigel Bird

Dinah Crystal OBE

Professor Geraint Howells

Paul Lowenstein QC

Gerard McDermott QC

Dr Margaret Mountford

Harriet Territt

This year's edition was generously funded by Jones Day.

The entire editorial team are sincerely grateful for Jones Day's support. This project would not have been possible without them.

Preface from Jones Day

The Law Society of England and Wales reports that, as at 31 July 2013, 158,644 individuals were admitted to the Roll of solicitors of England and Wales in 2013. Each one of those solicitors is subject to the first mandatory principle set down by the Solicitors Regulation Authority, to "uphold the rule of law and the proper administration of justice". The Bar Standards Board of England and Wales places the obligation similarly highly on its own list of objectives, as do many other international legal bodies and organisations. Jones Day has an equally strong commitment to the development of the rule of law in countries around the globe including, but not limited to, the 19 countries where our offices are located.

The range of subjects and complexity of the issues discussed in this volume of the Review reminds us that upholding the rule of law also means seeking to develop it through debate and practical application. In a fast-changing world, we must continually meet the challenge of new and difficult issues head on, whilst upholding the fundamental principles which underlie our legal systems. Reading the detailed contributions to this review also reinforces our strong belief that complexity is not always a vice, when it is accompanied by clarity of thought.

In that context, we are delighted to support this volume of the University of Manchester Review of Law, Crime & Ethics and to congratulate each and every contributor to the publication of this journal. We include in that group not just the writers, editorial team and Faculty members but every librarian, colleague, friend, flatmate or significant other whose support is an essential part of undertaking a project such as this one.

We particularly commend the writers and editors for their personal role in developing the rule of law with this volume

VIII

of the Review, and look forward to seeing their further contributions to their chosen field of law, business or study at some point in the future.

Harriet Territt
Partner
Jones Day

September 2014

Preface from the Head of the School of Law

It is my pleasure to welcome the third edition of this review. What started out as an idea from the student body has now become a feature of the Law School. It involves all disciplines of the School and students on all programmes as well as staff.

The papers show the keen interest of our community in exploring cutting edge issues, I hope all readers find something to reflect upon and many in our community are encouraged to contribute in the future.

The authors are to be congratulated, but so is the editorial team who have put in many hours of work. I step down as Head of School this summer, but am very pleased this review was launched during my tenure. I look forward to reading many more future volumes.

Geraint Howells

July 2014

Preface from the Editor-in-Chief

Editor-in-Chief's Foreword

At roughly the same time that I was appointed as Editor-in-Chief of this Review, Philip Burton, a fellow doctoral student within the School of Law at the University of Manchester as well as a fellow editor of this review, drew to my attention an article by Professor David Kennedy. The article – entitled ‘When Renewal Repeats: Thinking Against the Box’¹ – arrived in my inbox shortly thereafter. Being a mere English public lawyer interested in what public lawyers are generally interested in, I was, in truth, a little unsure of how much I would enjoy this article, which I was assured by Philip was a modern masterpiece of international law scholarship. This article, adding to the challenge of its rather alien content, stretched from page 335 to page 500 of the *New York University Journal of International Law and Politics*. I almost immediately closed the file and went on concerning myself with that which an English public lawyer ought to be properly concerned. I eventually got around to reading Professor Kennedy’s article on a long train journey and did so, I must admit, with the likelihood of returning to the book I was reading (about English public law) within a few minutes clear in my mind. The first thing that must be said of the article is trite to anybody who knows more about international law than I: Professor Kennedy’s article is a lucidly written observation of the nature of the discipline in which he operates and it undoubtedly deserves to be held in the high esteem which it is. However, it was the early passages of the article – which, probably to Philip’s dismay,

¹ David Kennedy, ‘When Renewal Repeats: Thinking Against the Box’ (2000) 32 *New York University Journal of International Law and Politics* 335.

had very little to do with international law - that caught my attention.

Professor Kennedy's article is a response to an invitation from the board of editors' at the *New York University Journal of International Law and Politics* - all of whom, as is often the case in the US, were students - to write a piece for their 'millennium issue'. This piece, they hoped, would set out both "new thinking" and Professor Kennedy's thoughts about the legal issues he expected to "shape the parameters of international law in the new millennium."² Professor Kennedy, aged 45 at the time, articulated his view on that invitation in the following terms:

"It may sound like a cop out, but I would be interested in the editors' response to their own initiative; in my experience students are often far closer to the next wave than older colleagues, and I have gotten a lot by picking up on the threads of their interests... there is something odd in a student editorial board asking people half again to twice their age what issues we think will become relevant in our careers - as if we were just starting out and they were situated too far along to be able to see that far forward."³

I thought this was a brilliant sentiment. One person I recalled who had made a similar observation is our School's very own Rodney Brazier, when writing in his foreword to the his *Constitutional Reform* book.⁴ The sentiment that young lawyers, untainted by the weariness of experience, can bring fresh ideas is one I have kept in mind whilst overseeing this project and one, to my mind at least, it is worthwhile for anyone who works within the walls of a University to remember. There is much learning by students, both undergraduates and postgraduates, that goes on within the walls of the University and it is, in many ways, a matter of

² *ibid* 335.

³ *ibid* 336.

⁴ Rodney Brazier, *Constitutional Reform: Reshaping The British Constitution* (3rd edn OUP; Oxford 2008) Foreword.

great shame how much of it never finds its way out. The *Review of Law, Crime and Ethics* strives to show the best of student scholarship within the Law School at the University of Manchester - which also incorporates the disciplines of criminology and medical ethics. This third issue of the Review, along with preceding issues, ultimately demonstrates all that is good about the community at the University of Manchester. Academics, students and practitioners coming together on a project that simultaneously demands and celebrates the hard work of those starting out in the law. Amongst these pages you will find articles on topics ranging from the proposed Common European Sales Law to the relationship between gender and crime. I hope you find them as enjoyable to read as I did.

I have reserved the most important issue until last: the settling of debts of gratitude. In producing this Review I have accrued many such debts to far too many people than I could possibly mention here. Those mentioned here are simply those who are due the most. Firstly, all the members of the editorial team have been committed, diligent and survived my, at times, extensive e-mailing. I hope many of them continue to work on the Review and I wish the best of luck to those moving onto new challenges. In particular, I would like to offer special thanks to Emile Abdul-Wahab, the Deputy Editor of the Review. Emile has been involved with the Review since its inception and much of what is best about the Review can be attributed to him. Dinah Crystal OBE has been a great source of support to the Review and, moreover, tolerant of my common unplanned excursions into her office in order to ask various questions. Thanks are also due to each member of the faculty at Manchester who advised on articles. This year has seen greater involvement from the faculty than ever before and the quality of the Review increases exponentially with their increased involvement. Kirsty Hawksworth, our web designer, has been a fantastic help in getting the review 'out' on the web - a much-needed expert for a near-luddite such as myself. Maureen Barlow's assistance throughout the year on various parts of the Review

is also genuinely appreciated - it is often criminally underestimated the cumulative contribution she make to the production of the Review. A huge thanks is also due to our sponsors, Jones Day. Without their support the Review would not be possible. In particular, it has been a pleasure to work with Diana Spoudeas and Keran Sandhu towards to successful competition of the Review.

Lastly, next year's Review will be edited by Amelia McGrath. Amelia has been an outstanding member of the editorial team for this year's issue and I look forward, with full confidence, to a fourth issue which surpasses, no doubt by some distance, this one.

Joe Tomlinson
Editor-in-Chief

August 2014

Contents

<i>PREFACE FROM JONES DAY</i>	VII
<i>PREFACE FROM THE HEAD OF THE SCHOOL OF LAW</i>	IX
<i>PREFACE FROM THE EDITOR-IN-CHIEF</i>	XI
<i>CONTENTS</i>	XV
BRITAIN'S PURSUIT OF A CODIFIED CONSTITUTION	1
<i>Arvinth Rai</i>	
THE 'CONSTITUTIONAL STRAIGHTJACKET' AND INADEQUATE RIGHTS-PROTECTION: A PROPOSAL FOR LIMITED JUDICIAL EMPOWERMENT IN THE HUMAN RIGHTS ACT	10
<i>Dalton Hale</i>	
ADDITIONAL MEMBER SYSTEM: THE FORGOTTEN ALTERNATIVE VOTE	23
<i>Philine Scheer</i>	
THE NATURE OF THE STATE	34
<i>Xian Liang Yuen</i>	
BEST INTERESTS CANNOT BE DIVORCED FROM FAMILY FAITH ..	44
<i>Stephanie Murphy</i>	
DIFFERENT CATEGORIES OF EMPLOYMENT STATUS: AN OPPORTUNITY FOR EMPLOYERS TO CUT COSTS OR ANOTHER HURDLE TO OVERCOME?	55
<i>Sophie Cooke</i>	
THE CESL PROPOSAL: PROGRESS OR MISSTEP?	67
<i>Sam Wardleworth</i>	
INCOHERENCE AND POLICY IN THE LAW OF CAUSATION: A CAUSE FOR CONCERN?	79
<i>Alex Sprake</i>	
GENDER, CRIME AND CRIMINOLOGY	88
<i>Chris Clarke</i>	

A COMPARATIVE ANALYSIS OF THE RNR BASED SEX OFFENDER TREATMENT PROGRAMMES	101
<i>Jessica Capello Mathews</i>	
A CRITIQUE OF ENLIGHTENED SHAREHOLDERS VALUE: REVISITING THE SHAREHOLDERS' PRIMACY THEORY	115
<i>Collins C. Ajibo *</i>	
FOOTBALL MUST PLAY BY THE SAME RULES AS EVERYONE ELSE - A COMMENT ON <i>BERG v BLACKBURN ROVERS</i> <i>FOOTBALL CLUB & ATHLETIC PLC</i>.....	139
<i>Philip Ebsworth</i>	
POWER v GREATER MANCHESTER POLICE AUTHORITY	144
<i>Amin Al-Astewani</i>	
THE QUEEN AND THE CONSTITUTION (AND ME).....	149
<i>Rodney Brazier</i>	

Britain's Pursuit of a Codified Constitution

Arvindh Rai¹

Abstract

This article attempts to flesh out some of the most important arguments for and against the codification of the British constitution. They are considered in the context of two distinct forms of codification, the first being the restatement of the UK's existing constitutional arrangements and the second being the reform of the British constitution. At the heart of these arguments lie assessments of parliamentary sovereignty and its compatibility with a codified constitution. This article also addresses the international peculiarity of the UK's position as the only European State without a codified constitution and its relevance to the pursuit of a codified constitution for the UK. It is concluded that, as things currently stand, there is no sufficiently cogent reason for the UK to pursue a codified constitution.

I. Introduction

The constitutional reform programme of the first New Labour Government of 1997-2010 gave rise to tidal waves of reform, some of which are still working their way through Britain's constantly morphing constitutional landscape.² The notion of a codified constitution for the UK appears to have ridden these waves, with both Labour and the Conservatives having indicated their commitment towards the production of a codified constitution³ and a British Bill of Rights respectively.⁴ The primary focus of this article is to assess

¹ Thanks are owed to Amin Al-Astewani and Professor Rodney Brazier for editorial support.

² Robert Hazell, 'The Continuing Dynamism of Constitutional Reform' [2007] ParlAff 3.

³ Gordon Brown in a speech on new politics at Centre Point: "And let me say to you today, that Labour's manifesto will include our commitment to charting a course to a written constitution."

⁴ David Cameron in a speech to the Centre for Policy Studies: "The Conservative Party, under my leadership, is determined to provide a hard-nosed defence of security and freedom. And I believe that the right way to do that is through a modern British Bill of Rights that also balances rights with responsibilities." This proposal was subject to wide criticism - see

whether the codification of the British constitution is a justifiable pursuit for the UK. This discussion will entail an outline of the main features of the British constitution, an assessment of its idiosyncratic nature and the implications of its, “essential unwrittenness”.⁵

At the outset, it should be noted that the codification of the UK constitution presents the analysis of two distinct forms.⁶ The first is the restatement of the UK’s current constitutional arrangements and fundamental constitutional principles in a single document.⁷ Conversely, the codification of the constitution could also be seen as an opportunity for extensive constitutional reform in the UK.⁸ This article shall discuss the significance, merits and limitations of both forms of codification.

II. Features of the British Constitution

The most fundamental feature of the uncodified British constitution is arguably the doctrine of parliamentary sovereignty. Traditionally, as defined by Dicey, it is Parliament’s constitutional right above all other institutions to legislate, amend or repeal any law of the UK.⁹ As Bogdanor noted, it is also the principal reason for the absence of a codified constitution.¹⁰ A codified constitution curtails Parliament’s legislative powers and is thus irreconcilable with the doctrine of parliamentary sovereignty.¹¹

Peter Munce, ‘The Conservative Party and Constitutional Reform: Revisiting the Conservative Dilemma through Cameron’s Bill of Rights’ [2014] *ParlAff* 80.

⁵ Michael Foley, *The Politics of the British Constitution* (Manchester University Press 1999) 186.

⁶ Nicholas Barber, ‘Against a Written Constitution’ [2008] *PL* 11.

⁷ *ibid* 12.

⁸ *ibid* 11-12; see also Vernon Bogdanor, *The New British Constitution* (Hart 2009) 14.

⁹ AV Dicey in Jeffrey Jowell and Dawn Oliver (eds), *The Changing Constitution* (7th edn OUP 2011) 40.

¹⁰ Bogdanor (n 8) 13; see also Vernon Bogdanor, Tarunabh Khaitan and Stefan Vogenauer, ‘Should Britain have a Written Constitution’ [2007] *Political Quarterly* 499, 501-502.

¹¹ This presupposes, foremost, that Parliament is sovereign. Criticism of this view is considered at pages 4-5.

The post-1997 constitutional reforms brought to light two other features of the British constitution, which are to be now considered. First, the disjointed manner and procedural failings of the reforms epitomised the British Parliament's cumulative, piecemeal approach towards constitutional matters.¹² Second, and more significantly, one of the greatest paradoxes of the reforms was that the inherent issue of precisely what the British constitution is precisely never came to the fore.¹³

To the likes of Griffith, the British constitution was purely political, consequent upon the balance of political power over time.¹⁴ Hennessey describes the constitution as a rich interplay of history and politics but makes no mention of the law.¹⁵ Johnson goes the furthest, lamenting the lack of legal or jurisprudential scrutiny as to what comprises the British constitution. He concludes that the British constitution "is not expressed or understood primarily in legal categories".¹⁶ Beatson however asserts that this is akin to sweeping aside the academic contributions of not only constitutional scholars but also the judiciary.¹⁷ Finally, Bogdanor plainly sums up the British constitution on the basis of one fundamental rule: "*What the Queen in Parliament enacts is law.*"¹⁸

¹² Rodney Brazier, 'How Near is a Written Constitution' [2001] 52 NILQ 1, 17.

¹³ Stephen Hockman QC and Vernon Bogdanor et al, 'Towards a Codified Constitution' (2010) 7(1) Justice Journal 74, 75.

¹⁴ John Griffith, 'The Political Constitution' [1979] 42 Modern Law Review 1; see also Jowell and Oliver (n 9) 3.

¹⁵ Peter Hennessey, *The Hidden Wiring: Understanding the British Constitution* (Victor Gollancz Ltd 1995) 30; see also Jack Beatson, 'Reforming an Unwritten Constitution' [2010] LQR 48, 49.

¹⁶ Nevil Johnson, *Reshaping the British Constitution* (Palgrave Macmillan 2004) 10; see also Beatson (n 15) 49.

¹⁷ Jack Beatson alludes to scholars varying from AV Dicey to Geoffrey Marshall as well as decisions including *Entick v Carrington* (1765) 95 ER 807 and *Council of Civil Service Unions v Minister for the Civil Service* [1985] AC 374 (HL) - Beatson (n 15) 49.

¹⁸ Bogdanor (n 8) 13.

III. Restatement of the UK's Constitutional Arrangements

The first form of codification that shall be considered entails the formalised restatement of the UK's existing constitutional arrangements. Bogdanor puts forward this enterprise for three main reasons. Bogdanor's first assertion revolves around the post-1997 constitutional reforms of the Labour government.¹⁹ He perceives these radical, far-reaching reforms as a precursory step towards the codification of the British constitution, seeing that many of our constitutional conventions have been legislated.²⁰ He thus perceives the codification process as a fitting consolidation of the piecemeal constitutional reforms since 1997.²¹

Secondly, Bogdanor cites the decline of the historical and conceptual reasons behind the absence of a codified British constitution. He perceives the post-1997 constitutional reforms as a "constitutional moment", which perhaps calls for the codification of the UK constitution.²² The Labour government in 2007 emphasised the need to determine what is meant to be "British" in the 21st century and indicated that a codified constitution may resolve this, hence further strengthening Bogdanor's belief that the UK may now be near a "constitutional moment".²³ Similarly, Bogdanor opines that the doctrine of parliamentary sovereignty, which stands as an obstacle to codifying the UK constitution, is in decline.²⁴ This would mean that parliamentary sovereignty is no longer a dominant principle in the UK and thus constitutes an opportunity to redefine our constitutional principles and, accordingly, enact a codified constitution.

¹⁹ Bogdanor (n 8) 215; see also Bogdanor, Khaitan and Vogenauer (n 9) 500.

²⁰ Hockman and Bogdanor (n 13) 75.

²¹ Vernon Bogdanor and Stefan Vogenauer, 'Enacting a British Constitution: Some Problems' [2008] PL 39.

²² Bogdanor (n 8) 215.

²³ HM Government, 'The Governance of Britain' (Green Paper, Cm 7170, 2007) 10, 40, 57, 62; see also Bogdanor (n 8) 216.

²⁴ Bogdanor (n 8) 215; Bogdanor and Vogenauer (n 21) 55-56.

Thirdly, Bogdanor laments the ambiguity behind what precisely the UK constitution entails. He likens it to the hypothetical scenario of someone joining a golf club only to be told that the rules are to be found in a myriad of committee and historical decisions and opines that this is precisely the position that UK citizens find themselves in today.²⁵

Conversely, Barber argues that the inherent ambiguity in parts of the UK constitution is indeed *beneficial*.²⁶ He highlights the ambiguous relationship between, firstly, the Westminster Parliament and the European Institutions and, secondly, the British Courts and the European Court of Justice. Barber points out that none of these legal and political institutions have unparalleled authority to determine the force of European laws and legislation.²⁷ He thus perceives this ambiguity as a beneficial compromise, i.e. an implicit “agreement to disagree”, and a framework in which conflicting views can co-exist.²⁸

Barber moves on to the *feasibility* of codifying the UK constitution and, specifically, the inherent difficulty between codifying what the constitution *is* and what it *ought* to be.²⁹ He opines that simply restating the constitution is impossible because it would inevitably incorporate some element of reform too.³⁰ Bogdanor and Oliver themselves acknowledge problems of *scope* (i.e., which constitutional principles and conventions are to be included in a codified constitution)³¹ and *authority* (i.e., who has the authority to formulate a codified constitution).³² All things considered,

²⁵ Bogdanor (n 8) 216; see also Hockman and Bogdanor (n 13) 74.

²⁶ Barber (n 6) 15.

²⁷ Barber (n 6) 16.

²⁸ Barber (n 6) 17.

²⁹ Barber (n 6) 12.

³⁰ Barber (n 6) 13.

³¹ Bogdanor (n 8) 216-217; Dawn Oliver, ‘Written Constitutions: Principles and Problems’ [1992] ParlAff 135, 146-148.

³² Bogdanor (n 8) 217.

Barber sees little benefit in the enterprise of codifying the existing constitution.³³

It is submitted that Barber's final point regarding there existing little benefit in codification (or rather, the lack of it) is inaccurate. The benefits of a codified British constitution are obvious. Codification provides a clear and coherent account of existing constitutional arrangements.³⁴ Even Dicey, who was against codification, acknowledged the enviable position of countries that have their constitution enshrined in an overarching document.³⁵ Baker perhaps states what Barber may have meant to say in a cogent and plausible manner: constitutional reform should take precedence over clarification of existing rules as the primary goal of codification.³⁶

Moving on to Barber's belief that ambiguity is beneficial in "*certain* circumstances" and "*parts* of constitutional law",³⁷ the question that begs an answer is: what about *parts* of constitutional law where ambiguity is not beneficial? Barber himself admits that there are uncertainties in the UK constitution that need to be resolved, for example the reform of the House of Lords.³⁸ Barber's argument is only applicable to certain parts of the UK constitution. This undermines an otherwise astute point, especially in the context where the *entire* British constitution is to be addressed.

We now address Bogdanor's assertion that declining parliamentary sovereignty constitutes an opportunity to codify the constitution. The irony of this argument lies in the fact that it undermines an alternative argument for the codification of the constitution: curtailing the unrestrained legislative powers of Parliament. To say that parliamentary

³³ Barber (n 6) 12.

³⁴ Bogdanor, Khaitan and Vogenauer (n 10) 499.

³⁵ AV Dicey, *An Introduction to the Study of the Law of the Constitution* (10th edn Macmillan Press 1959) 4.

³⁶ Sir John Baker QC, 'The Unwritten Constitution of the United Kingdom' (2013) 15(1) *Ecclesiastical Law Journal* 4, 19.

³⁷ Barber (n 6) 15.

³⁸ Barber (n 6) 18.

sovereignty is in decline implies a shift towards legal constitutionalism, which weakens the case for codification of the constitution, thus rendering Bogdanor's point self-contradictory.

The crux of the argument against codifying the existing British constitution revolves around its *feasibility*. The problems of scope and authority are admittedly daunting. Nonetheless, as Bogdanor noted, there is no reason to believe they are unconquerable.³⁹ However, it would be premature to make any conclusion at this stage without considering, perhaps, a more preferred attitude towards codification: reform of the UK constitution.

IV. Reform of the British Constitution

Lord Scarman advocated the reform of the UK constitution through codification for two reasons: the lack of existing checks and balances in the British constitution and the inadequate enforcement of rights of *all* members of British society.⁴⁰ Scarman lamented the loss of power by the state and the Lords to the Commons due to increased democratisation and the modern party system, which consequently handed the Commons supreme legislative powers.⁴¹ He further lamented the meagre protection of the rights of minority groups in the UK and attributed it to the enjoyment of power by a parliamentary majority.⁴² He therefore called for a codified constitution to shackle the legislative power of the Commons and secure the legal protection of the minorities.

Barber, conversely, argues that the shift towards legal constitutionalism has already occurred in the form of political and legal threats to parliamentary sovereignty (e.g., European Union membership and devolution), rendering Scarman's argument archaic.⁴³ Baker further suggests that

³⁹ Bogdanor (n 8) 230.

⁴⁰ Lord Scarman, 'Why Britain Needs a Written Constitution' (1993) 19 *Commw L Bull* 317, 319.

⁴¹ *ibid* 320-321.

⁴² *ibid* 321-322.

⁴³ Barber (n 6) 13-14.

judges are unsuitable for policy-making authority which a shift towards legal constitutionalism would entail.⁴⁴

It is pertinent to note that Scarman made his assertions before the post-1997 constitutional reforms. We have now seen a radical shift towards legal constitutionalism in the form of enacted legislation (e.g., Human Rights Act 1998). This is notwithstanding existing parliamentary constraints prior to 1997, e.g., European Union membership. The question that advocates of a codified constitution must thus answer is: are we in need of further reform, to the extent where codification of the constitution is the sole answer?⁴⁵ Indeed, there appears to be little need for such extensive reform. Furthermore, the courts have recently recognised potential “constitutional legislation”,⁴⁶ and even indicated limits to parliament sovereignty.⁴⁷ This casts the perception of Parliament’s unconstrained legislative powers in great doubt. Whilst we do not have the legal protection of a codified constitution, Parliament is sufficiently shackled to address Scarman’s concerns: the adequacy of checks and balances in the UK constitution, and the protection of minorities.

V. Conclusion: The Wider European Perspective

We now consider if being the only European State without a codified constitution is an adequate reason for embracing codification. In this respect, it is submitted that incremental constitutional development possibly leading to the codification of the UK constitution is a better alternative than the conscious desire to codify the constitution because other countries have a codified constitution. Furthermore, the legitimacy of any constitution is only as good as its content and the sacrosanct regard and respect accorded to it by its

⁴⁴ Baker (n 36) 19-20.

⁴⁵ Barber (n 6) 18.

⁴⁶ *Thorburn v Sunderland City Council* [2002] EWHC 195 (admin) [62]-[63] (Laws LJ).

⁴⁷ *R (Jackson) v Attorney General* [2005] UKHL 56 [102], [104], [159].

governors.⁴⁸ Whilst the Italian Constitutional Court's striking down of Berlusconi-backed legislation⁴⁹ highlights the legal protection a codified constitution provides, the subversion of the USSR Constitutions of 1936 and 1977 by the Communist Party⁵⁰ and the controversial 2004 Belarusian Referendum⁵¹ indicate that a codified constitution is not necessarily impenetrable. Therefore, the burden is on proponents of a codified constitution to not merely show the benefits of codification but to prove that the UK is in need of an entire constitutional overhaul that justifies codification.⁵² It is submitted that this contention is premature. As such, there is no sufficiently powerful reason for the UK to pursue a codified constitution as things currently stand.

⁴⁸ Harold Laski, *Liberty in the Modern State* (Faber & Faber 1937) 76; Harold Laski, *A Grammar of Politics* (Yale University Press 1925) 103-104.

⁴⁹ HadaMessia, 'Court strikes down key parts of Berlusconi law' (CNN, 13 January 2011) <<http://www.cnn.com.sg/2011/WORLD/europe/01/13/italy.berlusconi.immunity/index.html>> accessed 4 November 2013.

⁵⁰ Oliver (n 31) 139.

⁵¹ Carolynne Wheeler, 'Belarus face EU wrath for disputed poll result' (*The Guardian*, 19 October 2004) <<http://www.theguardian.com/world/2004/oct/19/eu.russia>> accessed 4 November 2013.

⁵² Barber (n 6) 11.

The ‘Constitutional Straightjacket’ and inadequate rights-protection: a proposal for limited judicial empowerment in the Human Rights Act

Dalton Hale

Abstract

The Human Rights Act 1998 (the “HRA”) does not sufficiently protect fundamental human rights. In order to do so the HRA must command the status of a “sacred charter”, but parliamentary sovereignty in the United Kingdom (“UK”) is a hindrance to adequate protection. The current scheme of rights-protection in the HRA compensates by means of two effective, but ultimately inadequate, forms of protection: a de facto ‘strike-down’ power, in the form of declarations of incompatibility; and ‘moral entrenchment’, which is founded on certain political disincentives that constrain outright repeal of the HRA. To remedy the shortfalls of these partial defences, the UK should follow the Canadian example and empower the judiciary with a limited, but de jure, strike down power. It is argued that such a re-distribution of power as between the judiciary and Parliament is not inconsistent with democracy, but admittedly is a limited abrogation of parliamentary sovereignty. It is however a necessary step if the rights contained within the HRA are to command the status of sacred rights beyond the reach of government; rights-protection is better served when the judiciary acts as the guardian of individual rights and as a check on the powers of government.

Introduction

This essay, written before December 2012,¹ aims to consider whether the Human Rights Act 1998 (the “HRA”) sufficiently protects fundamental human rights. It begins in

¹ And therefore before publication of: Commission on a Bill of Rights, *A UK Bill of Rights? The Choice Before Us* (2011-12, Volume I); Commission on a Bill of Rights, *A UK Bill of Rights? The Choice Before Us* (2011-12, Volume II). However, effort has been made to address some of the implications of this throughout.

Section I with the premise that in order to do so the HRA must command the status of a 'sacred charter', and describes the 'constitutional straightjacket' within which the HRA operates. Section II explains how the current scheme of rights-protection in the HRA compensates by means of two effective, but ultimately inadequate, forms of protection: a *de facto* 'strike-down' power and 'moral entrenchment'. Further, in Section III, it will challenge the view that a re-distribution of power as between the judiciary and Parliament is inconsistent with democracy and make a brief comparative analysis culminating with a proposal to empower judges by means of a limited, but *de jure*, strike-down power.

Section I

I. Entrenchment and the 'Constitutional Straightjacket'

What distinguishes a Bill of Rights from an ordinary rights-protecting statute is that it includes "some attempt to entrench their provisions as an attempt to establish them as a form of higher or fundamental law".² This usually manifests itself in two ways: (a) by a degree of entrenchment, and (b) by conferring power on the courts to override ordinary legislation.³

The effectiveness of this genus of rights-protection, it is submitted, stems from the public, judicial and parliamentary perception of the right-conferring instrument as a 'sacred charter' beyond the reach of government. Such sentiment can be seen from Hanna J in the Irish case *A-G v McBride*: "...the constitution is a sacred charter not to be lightly, vaguely, or equivocally tampered with.... The rights of the people should not be obscured by the facile pen of the draftsman".⁴

² Joseph Jaconelli, *Enacting a Bill of Rights* (Oxford University Press 1980).

³ Joint Committee on Human Rights, *A Bill of Rights for the UK?* (2007-08, HL 165-II, HC 150-II) Ev 137.

⁴ *A-G v McBride* [1928] 1 IR 451.

The central hindrance with which the United Kingdom (“UK”) must grapple, stems from Dicey’s propositions of parliamentary sovereignty: no Act of the Queen-in-Parliament may be held invalid by a court of law; no Parliament can bind its successors as to the form or content of subsequent legislation.⁵ Implicit within these propositions is the conclusion that if two Acts of Parliament conflict, the later statute will repeal the former; this doctrine has received judicial recognition.⁶ It is clear then, that most traditional forms of entrenchment, or the conferral of judicial strike-down powers, will plainly constitute a breach of parliamentary sovereignty. This is the constitutional straightjacket within which rights-protection in the UK must operate.

Section II

I. The HRA 1998: Core Provisions

The HRA was enacted in a way that does not sacrifice the supremacy of Parliament. By allowing the HRA to be repealed in the ordinary way, it has been given no special legal significance.⁷ This has been the source of much criticism. Of particular relevance is the view that the HRA is a merely “quasi-constitutional” piece of legislation that does not truly protect fundamental rights.⁸ Indeed it is argued, along these lines, that the provisions of the HRA do encompass a *degree* of protection; the question is whether this is enough.

⁵ Albert Venn Dicey, *An Introduction to the Study of the Law of the Constitution* (10th edn, Macmillan Education 1987).

⁶ *Vauxhall Estates v Liverpool Corporation*[1932] 1 KB 733; *Ellen Street Estates v Minister of Health*[1934] 1 KB 590.

⁷ In principle, a simple majority in the House of Commons will be enough to force legislation through the House of Lords via the Parliament Acts of 1911 and 1949. Royal Assent is granted by constitutional convention.

⁸ John Wadham, ‘Bringing Rights Half-Way Home’ (1997) *European Human Rights Law Review* 141.

The HRA is an enabling statute, which incorporates Articles 2 to 14 of the European Convention on Human Rights (ECHR) into UK domestic law. The core provisions pertinent to the scheme of rights-protection in the HRA are sections 2, 3, 4 and 19. Section 2 provides that when applying human rights law in a case, the judge must "take into account" the jurisprudence of the European Court of Human Rights (ECtHR). It is important to note that this does not mean that the ECtHR cases are *binding* on the judge.⁹ More significantly, section 3 contains an 'interpretative clause', so that in human rights related cases, judges must interpret the law compatibly with the ECHR "so far as it is possible to do so". In other words, the judge must not artificially construe or strain the meaning of the law in order to make it compatible with the HRA.¹⁰ Section 4 introduces "declarations of incompatibility" and Section 19 in brief, imposes an obligation on Ministers to 'flag-up' legislation enacted that is inconsistent with the scheme of rights-protection in the HRA. It is submitted that these provisions have developed two mechanisms of protection for the HRA which this essay dubs respectively, a *de facto* 'strike-down' power and 'moral entrenchment'. The existence and effectiveness of both are discussed in turn.

II. The *de facto* 'strike down' power: the interpretive clause and declarations of incompatibility

If it is not possible to interpret legislation so as to be compatible with the ECHR, the courts can strike down legislation - if it is subordinate legislation. However, they have no power to strike down primary legislation, for to do so would infringe parliamentary sovereignty. The compromise made by the HRA is that the courts may instead make a declaration of incompatibility.¹¹ Such declarations do

⁹ Paul Craig, *Administrative Law* (6th edn, Sweet & Maxwell 2008) 553.

¹⁰ It is worth pointing out here that this puts some strain on the system of precedent in the UK, as lower courts can effectively overrule higher courts on the basis of section 3.

¹¹ Human Rights Act 1998 (HRA 1998) s4.

not invalidate the law; they are designed to give a clear signal to Parliament that the particular legislation is incompatible with the HRA scheme of rights-protection and by implication should be brought into line.¹² The Commission on a Bill of Rights expressed its settled view that section 4 strikes a “sensible balance” between the “ultimate sovereignty of the UK Parliament” and the “duty of courts to declare and enforce the law”.¹³ Indeed, this mechanism, in combination with the ‘interpretive clause’, gives a significant degree of primacy to the ECHR and can be regarded as a *de facto* ‘strike-down’ power. Nonetheless, it might be contended that this mechanism is still too weak and a *de jure* judicial power to override incompatible primary legislation might be regarded as a more effective enforcement mechanism. There is some bite to this argument, in that Parliament is under no legal obligation to take any notice of declarations issued by the courts. A rebuttal to this concern is that within the context of British constitutionalism the difference between the power to strike down legislation and the power to declare it incompatible with the Convention “may be merely a technical one”.¹⁴ For, as Lord Hoffmann has observed, “if the courts make a declaration of incompatibility, the political pressure upon the government and Parliament to bring the law into line will be hard to resist”.¹⁵ Furthermore, if the Supreme Court issues a declaration of incompatibility and the government fails to respond, this is likely to trigger an application to the ECtHR. The applicant will then argue that their Convention rights have been violated and that the UK Government has nonetheless failed to remedy the violation. It is highly likely that the ECtHR would consequently find that Convention rights have been violated, resulting in the

¹² Peter Cane, *Administrative Law* (4th edn, Oxford University Press 2004) 47.

¹³ The Commission on a Bill of Rights, *A UK Bill of Rights? The Choice Before Us* (2011-12, Volume I) 35.

¹⁴ Douglas Vick, ‘The Human Rights Act and the British Constitution’ (2002) *Texas International Law Journal* 37.

¹⁵ Rt Hon Lord Hoffmann, ‘Human Rights and the House of Lords’ (1999) 62 *MLR* 159.

UK Government being subject to an obligation under international law to change the offending legislation in light of that ruling, adding even greater political pressure.¹⁶

A pertinent example of the relative success of the system can be found in *A v Secretary of State for the Home Department*,¹⁷ where 10 appellants were detained under section 23 of the Anti-terrorism, Crime and Security Act 2001, which provided for the indefinite detention without trial, and for the deportation of, non-British nationals suspected of terrorism. The judges in this case issued a declaration of incompatibility and three months later the Prevention of Terrorism Act 2005 repealed the provisions.¹⁸ Unfortunately the case also represents the shortfalls inherent within the current system. Indeed, Professor Rodney Brazier has described this as possibly an Act passed to try to get around the declaration.¹⁹ The new legislation, which introduces 'control orders', has been described by Amnesty International as "a grave threat to human rights and the rule of law in the UK",²⁰ which suggests that the effect of the section 4 declaration in this instance was completely undermined.

Another shortfall of the current system, as expressed by Professor Bogdanor in a 2006 lecture, is that it "might work in peaceful times, but in times of moral panic, Parliament and government might simply ignore a declaration of

¹⁶ Aileen Kavanagh, *Constitutional Review under the UK Human Rights Act* (Cambridge University Press 2009).

¹⁷ [2004] UKHL 56.

¹⁸ See also: *R (on the application of Wilkinson) v Inland Revenue Commissioners*[2003] EWCA Civ 814; *Blood and Tarbuck v Secretary of State for Health*(unreported); *R (on the application of Anderson) v Secretary of State for the Home Department*[2002] UKHL 46; *International Transport Roth GmbH v Secretary of State for the Home Department*[2002] EWCA Civ 158.

¹⁹ Email from Professor Rodney Brazier to author (16 November 2012).

²⁰ Amnesty International, *The Prevention of Terrorism Bill: A grave threat to human rights and the rule of law in the UK* (AI Index:EUR 45/005/2005, 28 February 2005)<<http://www.amnesty.org/en/library/info/EUR45/005/2005>> accessed 10 November 2012.

incompatibility”.²¹ Indeed, history has revealed that in times of crisis governments have “tended to overreact”.²² Thus, it is not unreasonable to conclude that this *de facto* ‘strike-down’ power, although generally effective, may fail to perform in certain circumstances.

Unfortunately, there is another, more substantial, flaw in the equilibrium between the protection of fundamental rights and the retention of parliamentary sovereignty. Even if a declaration is made and the offending legislation remedied, the individual injustice in the particular case before the court remains unaltered, primarily because legislation is not generally retrospective in application. To cite an example, in *Bellinger v Bellinger*,²³ Mrs Bellinger’s victory in obtaining a declaration that the Matrimonial Causes Act 1973 was a violation of her rights under Article 8 of the ECHR may have instigated the enactment of amending legislation,²⁴ but it had no effect whatsoever on her petition that her marriage to her husband was valid. Claimants may therefore have no incentive to bring their claims to court, as there is no opportunity to overturn the law in their favour. The result is a multiplicity of un-litigated claims, and thus numerous unresolved rights violations.

III. ‘Moral entrenchment’: the HRA and section 19

The second mechanism, ‘moral entrenchment’, rests on the *political disincentive* to enact legislation contrary to what are perceived to be universal human rights. Section 19 HRA forces the heads of incompatible government initiatives above the parapet to face the furore of the electorate. The section sets out the obligations incumbent on Parliament when enacting new legislation by providing that a Minister of the Crown in charge of a Bill in either House must, before

²¹ Vernon Bogdanor, ‘The Judges and the Constitution’ (Gresham College, May 2006) <<http://www.gresham.ac.uk/lectures-and-events/the-judges-and-the-constitution>> accessed 15 November 2012.

²² Tom Bingham, *The Rule of Law* (Penguin Books 2011).

²³ [2003] 2 AC 467.

²⁴ Gender Recognition Act 2004.

the second reading of the Bill, state that the provisions in the Bill are compatible with the ECHR.²⁵ If the Minister is unable to do so, a statement must be made to the effect that although incapable of making a statement of compatibility, the government nevertheless wishes the House to proceed with the Bill.²⁶ Not only does this ensure that Parliament always takes ECHR rights into consideration when enacting new legislation, but it also means that politicians will be reluctant to use section 19(1)(b), which requires Ministers to *expressly* state when there is inconsistency – something which will undoubtedly be unpopular with the electorate.²⁷

This kind of political constraint also entrenches the HRA in a more general sense. Although the HRA is an ordinary Act of Parliament, “it has already acquired such prestige in its short life that, politically speaking, repeal is out of the question”.²⁸ Indeed even before the enactment of the HRA, commentators recognised its political significance and the political disincentives which would make it unthinkable that “such a basic constitutional document could be overtly repealed”.²⁹ This is perhaps the most significant safeguard that the HRA has; it forms part of its ‘moral entrenchment’. Although David Cameron has proposed, and the Conservative Party will promise again in its 2015 manifesto, to repeal the HRA and replace it with a British Bill of Rights, such a move might require the abandonment of Britain’s international obligations to the ECHR, along with the vast body of case law that has firmly integrated the ECHR within English common law. Much care must be taken therefore, to avoid causing more harm than good. A sensible way forward

²⁵ s19(a) HRA 1998.

²⁶ s19(b) HRA 1998.

²⁷ In fact this has only been used once – when introducing the ill-fated House of Lords Reform Bill, Nick Clegg was unable to make a statement of compatibility under section 19(1)(a) HRA in respect of the provisions of that Bill relating to the laws on entitlement to vote at House of Lords elections, including the rules which prevent prisoners serving sentences from voting. Mr Clegg nevertheless wished to proceed with the Bill.

²⁸ Cane, *Administrative Law* (n 12).

²⁹ Jaconelli, *Enacting a Bill of Rights* (n 2).

might be for any replacement statute to reflect the substance of Convention rights – there is nothing preventing additional ‘home-grown’ rights then being integrated. It is crucial, however, that no *less* protection be provided under a new Bill. This suggestion is consistent with the conclusions of the Commission on a Bill of Rights.³⁰

For these reasons, it is not outright repeal that threatens the HRA³¹ but rather the potential for its “gradual and surreptitious erosion”³² by the implied repeal of its provisions. Section 19 may be incapable of militating against this kind of threat; implied repeals, insufficiently explained or overlooked by the Minister in charge of the Bill will likely pass by the electorate unnoticed.

The upshot of all of this is that in the end the rights contained within the HRA may not always be “rights beyond [the] reach of government”.³³ The importance of this point cannot be understated. According to Forest J of the Court of New Brunswick, the key psychological difference between, for example, the 1960 Canadian Bill of Rights³⁴ and the 1982 Canadian Charter³⁵ was the way that the latter was perceived by the judiciary and parliamentarians alike: it was seen to be beyond the reach of government.

³⁰ The Commission on a Bill of Rights, *A UK Bill of Rights? The Choice Before Us* (2011-12, Volume I) 30.

³¹ Although, of course, it could be. David Cameron’s proposals tell us that much, but he still aims to replace it with something of at least equivalent virtue. To make rights protection worse or indeed repeal and not replace the HRA would be very unlikely indeed.

³² Jaconelli, *Enacting a Bill of Rights*(n 2).

³³ Gérard La Forrest, ‘The Canadian Charter of Rights and Freedoms: An Overview’ (1983) 61 *Canadian Bar Review* 19.

³⁴ This endeavour was particularly unsuccessful. See Gerry Ferguson, ‘The Impact of an Entrenched Bill of Rights: The Canadian Experience’ [1990] *Monash University Law Review* 12.

³⁵ Canadian Charter of Fundamental Rights and Freedoms 1982.

Section III

1. The Judiciary, Parliament and Democracy

Many who are loath to accept entrenchment of the HRA point out that such a change would amount to a constitutional re-distribution of power from the legislature to the judiciary. This, they contest, is not in the interests of democracy, as it is Parliament who is elected by the majority to govern, not the judiciary.³⁶

But if judges were to have a more significant role in human rights protection, would this necessarily be a bad thing? Would it even be inconsistent with democracy? The right of the majority to rule according to Majoritarian theorists is the embodiment of the fundamental value of democracy,³⁷ but Dworkin asserts that Majoritarian procedures are but one contributing factor towards the true value of democracy: equality. As such, Majoritarian procedures can be flouted in the interests of respecting the democratic conditions of equality.³⁸ One democratic condition could be that *every* citizen is entitled to the same rights *equally*, and that these should be firmly secured within the constitution.³⁹ Thus, if judges could invalidate incompatible legislation and as a result, provide that any reform that has the implicit support of the majority alone⁴⁰ cannot bypass an entrenched Bill of Rights. The re-distribution of power is not inconsistent with, but *further*s, democracy. It is also worth noting that the judges are subject

³⁶ See discussion in Keith Ewing, 'The Human Rights Act and Parliamentary Democracy' (1999) 62 MLR 79.

³⁷ Jeremy Waldron, *Law and Disagreement* (Oxford University Press 1999).

³⁸ Ronald Dworkin, *Freedom's Law: The Moral Reading of the American Constitution* (Oxford University Press 1996).

³⁹ For the moral case for this, based on social contract theory, see John Rawls, *A Theory of Justice* (Harvard University Press 2003).

⁴⁰ At the exclusion of everyone that forms part of the minority. The dangers of Majoritarian democracy are particularly acute where there are indigenous minorities, see discussion in Geoffrey William George Leane, 'Enacting Bills of Rights: Canada and the Curious Case of New Zealand's "Thin" Democracy' (2004) 26 Human Rights Quarterly 152.

to a *democratically enacted* list of rights, rather than being entirely free to develop the common law unconstrained by such a text.⁴¹ This acts as a sufficient limitation on judicial power.

In any event, it may be questioned whether it is truly in the interests of democracy that the state is the omnipotent arbiter of its citizens' fundamental rights, when it is protection from state interference of those rights that a Bill of Rights is designed to protect. Perhaps it is best to assign roles to both Parliament and the Judiciary in the protection of ECHR rights, rather than putting rights-protection exclusively in the hands of one institution - it is worth remembering that it is only in the event of litigation alleging a rights-violation that the courts can intervene in order to articulate what those rights require in the context of a particular case; as such, the courts provide a "corrective mechanism" which forms a *part* of the system of rights-protection but "they are not, cannot and should not be the whole of that system".⁴²

II. Proposal for additional protection: the Canadian experience

The most significant aspect of the Canadian Charter is the fact that it puts into effect a "realignment of the balance of power" as between the legislature and the judiciary⁴³ - the courts have been given the power to determine whether government can legitimately pursue its legislative policies, and may accordingly invalidate legislation that is incompatible with the Charter.

Crucially for current purposes, the Canadian Parliament retains part of its legislative sovereignty by virtue of section 33 of the Charter. This allows the Canadian Parliament to enact any law that expressly declares that the enactment operates "notwithstanding" a potential violation of the Charter.

⁴¹ Aileen Kavanagh, *Constitutional Review under the UK Human Rights Act* (Cambridge University Press 2009).

⁴² *ibid.*

⁴³ Gerry Ferguson, "The Impact of an Entrenched Bill of Rights: The Canadian Experience"[1990] *Monash University Law Review* 12.

However, its use may be limited by the practicalities of politics,⁴⁴ as politicians will generally be reluctant to use it since they must expressly admit that their proposed legislation is in violation of fundamental rights and freedoms, something which would undoubtedly be unpopular with the electorate.

This seems to be a particularly robust system of rights protection. Parliamentary sovereignty is left largely intact by including the possibility of express abrogation from the Canadian Charter, while the political reluctance to do so, in combination with the judicial power to strike down certain legislation, gives such rights the requisite "sacred" status. This is broadly similar the current situation in the UK, with section 19(1)(b)⁴⁵ acting as the UK counterpart to section 33. Thus, the key reform proposal is that the UK courts should be able to declare any legislation unaccompanied by a section 19 statement invalid.⁴⁶ The attraction of this proposal is that parliamentary sovereignty would ultimately be retained by virtue of section 19(1)(b) and government would only use it where absolutely necessary, through fear of public outrage. Yet where government fails to acknowledge the incompatibility, the courts would be permitted to conclude that Parliament has overlooked the human rights implications of the enactment and can therefore treat it as invalid. It is not a strike-down power in its traditional sense; it is more subtle and heavily circumscribed as it *only* applies when government does not comply with the basic requirements of section 19. One may ask what this proposal adds to the UK system, and may regard such a *de jure* judicial power as merely a declaration of incompatibility "on steroids" - but its significance is that it would militate against

⁴⁴ Or, 'moral entrenchment'.

⁴⁵ Human Rights Act 1998.

⁴⁶ There was no appetite in the Commission report for conferring a power on the judges to strike down legislation. But there was also no consideration of this more *limited* kind of power - whereby Parliament can retain ultimate power. See The Commission on a Bill of Rights, *A UK Bill of Rights? The Choice Before Us* (2011-12, Volume I) 35.

the gradual and surreptitious erosion of the HRA, compel Parliament to reform the law in accordance with section 19,⁴⁷ and also allow the courts to remedy *individual injustices*.

Conclusion

Although there exist two substantial protection mechanisms in the HRA, there remain significant shortfalls in the current system. The *de facto* 'strike-down' power is often undermined when Parliament 'legislates around' declarations of incompatibility and further, such declarations do not allow individual injustices to be rectified. Further, 'moral entrenchment', whilst militating against outright repeal of the HRA may fall short of safeguarding the HRA from gradual and surreptitious implied repeals which do not attract the critical eye of the electorate. By taking inspiration from the 'Canadian experience', and empowering the judiciary with a limited, but *de jure*, strike-down power, it can be ensured that: a) the HRA is protected against gradual, implied repeals that may go unnoticed; b) where an implied repeal has been exposed by the court the government is compelled to adequately deal with it, and follow the section 19 mechanism correctly, rather than simply 'legislating around' a declaration of incompatibility; c) individual injustices can be rectified; and d) the democratic conditions of equality are better respected. The proposal constitutes a limited⁴⁸ abrogation of parliamentary sovereignty but is a necessary step if the rights contained within the HRA are to command the status of sacred rights beyond the reach of elective dictatorship; rights-protection is better served when the judiciary acts as the guardian of individual rights and as a check on the powers of government.

⁴⁷ Unless it chooses not to replace the invalidated legislation at all, in which case there will be a victory for human rights.

⁴⁸ Limited because Parliament can still expressly legislate contrary to the HRA if it uses section 19(1)(b) - it is only when this process is not used that this judicial power will be engaged.

Additional Member System: The Forgotten Alternative Vote

Philine Scheer

Abstract

Germany is frequently lauded as an example both economically and socially, with its stable trade surplus, progressive industrial relations and a strong sense of civic duty. There is certainly much to be admired. However, one area which British politicians seem unwilling to seek inspiration from is the role a fair electoral system plays in creating these conditions.

The British First-Past-the-Post (FPTP) System has received substantial criticism in recent years which gave rise to a referendum in 2011.¹ The referendum revealed public discontent with the present voting system, but also that the Alternative Vote was not the answer.² Modelled on the German constitution, I will propose a form of proportional representation, the Additional Member System. In part I, I will highlight the weaknesses of the UK voting system, before demonstrating in part II why a purely party list based system of proportional representation is not desirable either. In part III, I will endorse the German Additional Member System (AMS) as a compromise and remedy for FPTP's shortfalls.

I. The First-Past-the-Post System

In this section I will discuss how FPTP ensures a close connection between voters and the electorate, and its historical effectiveness for small constituencies. In today's multicultural society, however, it is reductive, unrepresentative and leads to overly powerful governments.

¹ Electoral Commission (2011) <http://www.electoralcommission.org.uk/find-information-by-subject/elections-and-referendums/past-elections-and-referendums/referendums/2011-UK-referendum-on-the-voting-system-used-to-elect-MPs> (accessed 14/04/2014).

² Electoral Commission (2011) Referendum on the voting system for UK parliamentary elections: Report on the May 2011 referendum.

Under FPTP voters mark an 'X' against a single candidate on the ballot paper.³ In order to win, a candidate must collect more votes than any other candidate in the constituency. Hence, a plurality of votes is needed, not a majority. The very name of the voting system is therefore misleading, as it appears to require the winner to meet a certain threshold. Those who support FPTP emphasise that it is a simple voting system that rarely results in hung parliaments and produces one-member constituencies with a potentially strong link between Members of Parliaments (MPs) and constituents.

However, FPTP regularly produces disproportionate results at general elections. Governments can be formed with a comparatively small proportion of the popular vote and can even form a majority with less overall votes than the opposition. Until the last election, all governments since World War II have been formed by single parties that gained less than 50 per cent of the vote⁴. On two occasions, in 1951 and 1974, the dominant party in the House of Commons has earned fewer votes than the runner up. In the general election of 2001 Labour won 64 per cent of the seats in Great Britain on 42 per cent of the votes cast. The turnout at the election was 59 per cent, so the vote for Labour represented only 25 per cent of the registered voters.⁵ This raises questions about the legitimacy of government and has discouraging effects on the electorate. In the following paragraphs I will discuss how these discrepancies, demonstrating a relatively low effective vote, result in 'wasted votes', entry barriers to parliament for independent candidates, smaller parties, and a low election turnout.

On 6 April 2010, the Electoral Reform Society estimated that of the 650 constituencies, 382 (59%) were 'safe seats' for

³ Martin Linton and Mary Southcott, *Making Votes Count: The Case for Electoral Reform* (Profile Books Ltd 1998) 134.

⁴ Dawn Oliver, *Constitutional Reform in the UK* (Oxford University Press 2003) 138.

⁵ Electoral Commission *Election 2001. The Official Results*. (London: Politico's, 2001) 220.

individual parties.⁶ Constituents who are not supporters of the prevailing party in their area have virtually no chance of making their vote 'count'. Likewise, voters backing the dominant candidate might feel that their ballot papers merely act as increasing an already large majority. Shifting opinion can also be easily exaggerated through considerable swings in seats from one major party to another. Landslide majorities might not closely reflect choices made by the electorate.

Once elected with an extensive majority, the party in power may neglect opinion in Parliament and among the electorate. Smaller parties are generally precluded from Parliament aside from those with a strong footing in a particular region such as Plaid Cymru in Wales or, lately, the Green Party in Brighton. Niche parties with a fairly even geographic distribution of supporters, on the other hand, have dim prospects of entering the House of Commons. Independents tend to have little hope of election; prominent exceptions include George Galloway and Dr Richard Taylor. This limits the variety of voices in Parliament and silences those who do not endorse one of the major parties.

In constituencies where the majority of the electorate did not support their MP in the election, voters can feel unrepresented in Parliament and a sense of alienation and injustice amongst disenfranchised sections of the electorate emerges.⁷ This disempowerment of voters reflects badly on the turnout. Elections in Britain compare unfavourably with other EU countries: the general election in 2010 had an unusually high turnout of 65.3 per cent,⁸ whereas the last German national election in 2009 noted a turnout of 70.3

⁶ Electoral Reform Society (2010) <<http://www.electoral-reform.org.uk/safe-seats/>> accessed 17 April 2013.

⁷ Dawn Oliver, *Constitutional Reform in the UK* (n 4) 133.

⁸ John Walton 'Election 2010: Turnout mapped' *BBC News* (10 May 2010) <http://news.bbc.co.uk/1/hi/uk_politics/election_2010/8672976.stm> accessed 16 April 2013.

per cent.⁹ This outcome reflects previous years, when British participation tended to be even lower and German turnout higher. Although this may be attributed to factors such as trust in government, the competition between parties, media coverage and apathy in respect of voting as civic duty,¹⁰ the peculiarities of the British FPTP itself appear to negatively impact voter turnout.¹¹ Other forms of proportional representation do not have the same setbacks.

II. Proportional Representation

As the term suggests, 'Proportional Representation' (PR) means that the purpose of the electoral system is to ensure that each party wins a proportion of seats in the elected body that closely reflects their popularity with voters, enhancing fairness between the parties.¹² Many Western European countries, for instance, adopt party list systems, whereby voters select one party list with a given order of candidates (closed-list system) or rank the listed candidates according to their preference (open-list system).

A pure form of PR, however, would be unlikely to enjoy wide support in the UK. Party list systems generally require dividing the country into bigger regions (or even the entire country, as in the case of Israel and the Netherlands), as opposed to the relatively small constituencies in the UK. Each of these constituencies is provided with a number of seats usually depending on its population and the parties submit lists with their candidates. Although independents are allowed to run without party endorsement, again, in practice they have little chance of being elected.

⁹ Bildungszentrale für politische Bildung (2009) <<http://www.bpb.de/nachschlagen/zahlen-und-fakten/wahlen-in-deutschland/55588/nach-bundeslaendern>> accessed 14 April 2013.

¹⁰ Catherine Bromley, John Curtice and Ben Seyd, 'Political engagement, trust and constitutional reform' in A Park et al (eds) *British Social Attitudes: The 18th Report* (Sage, 2001).

¹¹ Electoral Commission *Election 2001. The Official Results*. (Politico's 2001) 13.

¹² Dawn Oliver, *Constitutional Reform in the UK* (n 4) 133.

One consequence of this feature is that the party list system increases the distance between the elected representative and the constituency. Unlike FPTP, there is no easily identifiable MP for one region with the same direct relationship to the electorate. This would abolish the concept of the 'Constituency MP', which is deeply embedded in the political culture of Great Britain¹³ and perceived as an important link between Westminster and local communities.

Furthermore, closed-list systems transfer considerable patronage and power to the party leadership, who select and rank their candidates. This system limits voters' choices and allows unpopular candidates to merit from the party label. Although open and partially open lists allow for a greater choice and more personal vote, their effectiveness depends on the size of the list. A long list with about twenty names requires the voter to be familiar with all the candidates in order to express a preference and would tempt the party leadership to recommend candidates, which would undermine the objective of the exercise.

In sum, pure party list systems enable a highly effective vote by placing at least one, and potentially several, representatives of one's preferred party in Parliament. This would decrease the amount of "wasted votes"¹⁴ and potentially encourage turnout; however, it would sever the direct relationship between the voter and the MP. Furthermore, this variant of PR would still be unfair to independents and potentially result in a multiplication of parties causing a ruptured parliament.¹⁵

¹³ John Curtice and Ben Seyd, *Wise after the Event? Attitudes to Voting Reform Following the 1999 Scottish and Welsh Elections* (Constitution Unit 2000).

¹⁴ Stephen Fisher 'The Wasted Vote Thesis: West German Evidence' (1973) 5(2) *Comparative Politics* 293.

¹⁵ John Curtice, Stephen Fisher and Laurence Lessard-Phillips, 'Proportional Representation & the Disappearing Voter' in *British Social Attitudes, 23rd Report* (Sage Publications 2007) 129.

III. The German Additional Member System

The Additional Member System (AMS) currently used for Bundestag (national Parliament) elections in Germany attempts to marry single member constituencies with proportional representation, and provides both a remedy for the weaknesses and a combination of the strengths of both FPTP and PR. The German example appears useful, as Germany is a Western European democracy of comparable size and similar socio-economic structure to the UK. Furthermore, the AMS has been the electoral system of choice for the Scottish Parliament, the Welsh Assembly and the Greater London Assembly, showing its compatibility with British political culture.

The growing use of the AMS across the globe¹⁶ supports the claim that this system provides the “best of both worlds”¹⁷: locally, it ensures a direct relationship between constituents and their representatives, while nationally it ensures parties enjoy a share of power which is proportionate to their share of the vote. The nature of the electoral system combines the different perspectives of the negotiating parties in post-war Germany. The Western allies endorsed a majoritarian system, which featured their own MP-constituent nexus. The Länder (regional) governments in contrast, which had to design their own electoral laws, unanimously advocated proportional representation.¹⁸ These conflicting principles are not just represented in the general framework of the German AMS, but also in the specifications of the system.

Under AMS, each voter is awarded two votes – one for a candidate, and one for a party. In the style of FPTP, each constituency returns a single candidate. The votes for the

¹⁶ UK, New Zealand, Mexico, Bolivia and Lesotho.

¹⁷ Robert Johns, *AMS in Germany - and in Britain* (Democratic Audit) <www.democraticaudit.com/download/C-German-AMS.doc> accessed 14 April 2013.

¹⁸ Eckhard Jesse, ‘The electoral system: More continuity than change’ in Ludger Helms (ed) *Institutions and Institutional Change in the Federal Republic of Germany* (Macmillan 2000).

party list are then divided amongst the party's candidates and added, in addition to their constituency seats, to their number of seats in Parliament. These are the "additional members".¹⁹

In Germany this translates into two categories of seats in the Bundestag: 50 per cent of the MPs are selected in single-member constituencies; the remaining 50 per cent are derived from party lists. Similarly, electors are awarded two votes every four years: the first (Erststimme) for their preferred constituency representative, the second (Zweitstimme) for a regional party list. The list element serves to counteract the disproportionality inherent in the constituency component.

In the paragraphs below I will outline how the German AMS allows for a wider variety of voices in Parliament without providing a stage for radical movements. Furthermore, I will demonstrate that the AMS increases gender diversity, and creates stable and efficient governments.

Before the 2011 referendum, proponents of FPTP regularly argued that a move away from our existing voting system would carry the danger of inviting extremist parties into the legislature. Under Germany's AMS, a party has to gain at least 5 per cent of the popular vote to access the Bundestag, even if it is unsuccessful in securing any directly contested seats. This feature of the German electoral system was introduced to eliminate fringe movements, which spawned during the Weimar Republic, whilst enabling inclusion of smaller parties. The pro-business Free Democrats, the Greens and the Left Party usually meet this threshold.²⁰ However, when the NPD, an extremist right wing party, loomed in 1969 it was obstructed from entering the

¹⁹ Electoral Reform Society (2010) <<http://www.electoral-reform.org.uk/additional-member-system/>> accessed 14 April 2013.

²⁰ Leon Mangasarian, 'How Germany's Election System Works: What to Watch for Tomorrow' *Bloomberg* (25 September 2009) <http://www.bloomberg.com/apps/news?pid=newsarchive&sid=aBSHoe_o5NX8> accessed 14 April 2013.

Bundestag by not meeting the minimum threshold. Even if extremist parties should overcome the 5 per cent hurdle, this can serve to galvanise the centre parties and force them to address extremist trends with policy, since fringe movements are generally a symptom of deeper entrenched problems in society, such as youth unemployment or increasing class divides.

A major concern in the UK in recent years has been the disproportionately low number of women that have been elected into office.²¹ After the general election in 2010, only 22 per cent of MPs were women.²² This is because parties face a dilemma in the selection process and must deliberate between selecting the candidate they think potential supporters are more inclined to support – traditionally the middle-aged white male – or to be unbiased and risk losing the constituency. Consequently, the marginal representation of women in elected bodies may result in a negative effect on turnout, because the concerns and perspectives of female voters will not be adequately addressed.²³ In Germany, female representation in Parliament totalled 32.8 per cent after the last election in 2009.²⁴ Although this still falls short of equality, it is 10 per cent higher than in Britain and shows that AMS with its list component facilitates the selection and election of more women. This is supported by evidence from the Scottish, Welsh and London Assemblies, which have achieved better women's representation compared to Westminster.²⁵

²¹ John Morrison, *Reforming Britain. New Labour, New Constitution?* (Reuters 2001).

²² House of Commons Information Office Factsheet M4, 'Women in the House of Commons' (House of Commons 2010).

²³ Fiona MacTaggart, *Women in Parliament: Their Contribution to Labour's First 1000 Days* (Fabian Society 2000).

²⁴ Bildungszentrale fuer politische Bildung (2011)
<http://www.bpb.de/gesellschaft/gender/frauen-in-deutschland/49418/frauenanteil-im-deutschen-bundestag> (accessed 15 April 2013).

²⁵ *The Review of Voting Systems*, Cm. 7304 (2008) <<http://www.officialdocuments.gov.uk/document/cm73/7304/7304.asp>> accessed 14 April 2013.

The most frequent argument against AMS is that it will produce weak and unstable coalition governments, which outweighs discrimination against small parties and their supporters in the UK's electoral system. As with the current government, voters would regularly have to accept a combination of parties they did not vote for and a policy patchwork that was not proposed to them. However, this does not need to be a burden; in the Federal Republic of Germany coalitions have been the norm for most of the post-World War II era. They invariably included at least one of the two main parties (the SPD or the CDU/CSU) and the FDP (the small liberal democratic party) or the Green Party. Prior to each election, the FDP and the Greens respectively stated which other party they would form a coalition with and set out their conditions for joining this party. Whilst the SPD and the CDU/CSU name 'chancellor-candidates', the smaller parties generally do not propose a candidate, as it is very unlikely for them to be elected. Alternatively, they name one or two politicians who will lead their parties' campaign and will usually receive one of the most important ministerial posts if the party gets elected.

If both the small and main UK parties affirmed their preferred coalition partner, or the terms on which they would join or include another party, then the British public would know what they were deciding for (or against) when making their cross on the ballot paper.

The argument that one-party governments are more desirable is often based on the perception that they are strong and effective, compared to weak and indecisive coalition governments.²⁶ However, a single-party government is not inherently either strong or effective. From 1974 to 1979 the Labour government did not accomplish their policies of defeating inflation and rebuilding the economy; however they only enjoyed a narrow majority after the second election of 1974. This government was not only ineffective but also

²⁶Michael Dummett, *Principles of Electoral Reform*(Oxford University Press 1997) 178.

indecisive, reversing its policies of high public expenditure and adopting austerity measures.²⁷ Interestingly, the subsequent Conservative government demonstrates the dangers of the other extreme. In terms of adopting legislation and presenting strong leadership, the ‘Iron Lady’s’ governments certainly succeeded. But what if we do not understand ‘strong government’ in terms of determination and resistance to pressure and criticism? If we acknowledge the Thatcher government’s close proximity to authoritarianism²⁸ and lack of accountability, this era might quickly turn it into a case for coalition governments. A government that possesses a secure majority under the current system tends to be insensitive to concerns from inside Parliament. Only its own backbenchers might be able to have some impact but the concessions to prevent a backbench revolt are likely to be marginal. The Thatcher government, for instance, was able to adopt the poll tax, privatise the water industry and execute other unpopular policies during their legislative period.

Despite this, some people still maintain that a coalition would result in patchwork policies and instability but this is less problematic than it appears. In reality, voters do not endorse every single policy of their preferred party and might even welcome a considerable number of points in the programme of the rival party.²⁹ This also helps to eradicate unpopular and controversial policies that are mixed into a mainstream party’s manifesto.

Against the current coalition government’s initial difficulties and deep controversies about the introduction of tuition fees, it is not less effectively functioning than previous governments. Under the Cameron/Clegg leadership, earnings

²⁷ Michael Artis et al ‘Social democracy in hard times: the economic record of the Labour government 1974-1979’ (1992) 3(1) *Twentieth Century British Century* 33.

²⁸ Bob Jessop et al ‘Authoritarian populism, two nations and Thatcherism’ (1984) 147 *New Left Review* 32.

²⁹ Richard Rose, *The Problem of Party Government* (Penguin Books, 1976) 305.

have caught up with inflation and unemployment fell below seven per cent for the first time in five years, a remarkable achievement for a “weak” coalition government.³⁰

Similarly, it would be unjustified to characterise the coalition governments derived from AMS as unstable. Between 1949 and 2013, Germany saw only eight Chancellors, who were (with the potential exception of Kurt-Georg Kiesinger) powerful statesmen in their own ways.³¹ Every government or Parliament existed at least three years whilst in Britain there were 14 Prime Ministers, including three parliaments which did not survive even two years.

IV. Conclusion

Rather too much than too little stability then seems to be a weakness of the German system. Yet the 1989 election completely replaced the government as a result of the electorate’s dissatisfaction just as it happened in Britain in 1997, however, without the exaggerated discrepancy, which is characteristic for FPTP. More importantly, the German electoral system has not only contributed to stable but also, altogether, excellent government. In terms of economic success, a liberal and tolerant regime at home and a peaceful and responsible foreign policy, it is hard to find a country of similar size and influence, which has performed better or as well over the last 60 years. Particularly in view of Germany’s preceding history, it is a great achievement. It would be unreasonable to credit all this to the Additional Member System, however, there is powerful evidence to suggest that adopting a similar system in the UK would carry many of its benefits.

³⁰ BBC, ‘ONS: After six years, wage rises match inflation’ (16 April 2014) <<http://www.bbc.co.uk/news/business-27047966>> accessed 16 April 2014.

³¹ Report of the Independent Commission on the Voting System (The Jenkins Report) Cm 4090 (October 1998) Ch 4.

The Nature of the State

Xian Liang Yuen

Abstract

The determination of statehood is structured around legal and factual approaches. Rather than arguing for the sole domination of either approach, this article aims to show that attainment of statehood is a process involving both law and fact. This article shall examine the Montevideo Convention and other criteria perceived to be essential for gaining recognition in order to underline how law and facts lead to state recognition.

I. Introduction

When ascertaining whether an entity is a state there are two main methods: statehood can be determined by reference to facts¹ or by law.² It is submitted that neither of these two approaches can single-handedly account for explaining statehood in international law. This paper aims to study these methods in order to show that the nature of the state is a complex amalgam of legal and factual implications, and that the complete disregard of either one will result in a circular understanding of statehood. Much of the discussion of statehood revolves around the idea of state recognition; both the conditions for granting recognition and the consequences to be attached from it are controversial. The article begins with a discussion of the prevalent theories of recognition followed by an examination of the Montevideo Convention and other criteria for recognition.

*The author would like to thank Professor Jean d'Aspremont, Philip Burton and Janus Lim for their helpful comments. Any errors found in this article are made by the author only.

¹ Formation of a new state is largely a question of fact rather than law: Lassa Oppenheim, *International Law, Vol 1* (1st edn) 264; Law is secondary to factual decisions: Carl Schmitt, *Political Theory: Four Chapters on the Concept of Sovereignty* (University of Chicago Press 1985) 31.

² Hans Kelsen, 'Der soziologische und der juristische Staatsbegriff - The Sociological and Legal Concept of the State' (1928) Scientia Verlag 1.

II. Constitutive and Declaratory

There are two main theories regarding the consequences attached to state recognition: the constitutive and declaratory theories.³ The declaratory theory places emphasis on the pre-existing state of affairs. Recognition is understood as a willingness to enjoy diplomatic relations rather than a statement about the nature of the entity in question. However, facts alone cannot speak for themselves and must be interpreted through legal processes before legal consequences can arise. As Kelsen noted, a purely declaratory view will “confuse” facts with law.⁴

The constitutive theory holds that an entity attains statehood only after other states on the international plane decide to recognise it.⁵ However this stance opens the door for subject identification to become a political process. Since states are free to grant recognition as they see fit based on whatever reasons or incentives they have, the law seems to have no control over the determination of the primary legal subjects. However as noted above, the purely legal approach will not fly. The nature of the state has to be studied wearing the lens of both theories if one is committed to understanding statehood.

III. The Montevideo Convention

This section covers two criteria of the Montevideo Convention (the ‘Convention’) in order to argue that, despite having this legal conception of statehood, the nature of the state is in fact an amalgam of the legal and factual elements. Despite its many critics,⁶ the Convention is commonly used

³ Martin Dixon, R McCorquodale and S Williams, *Cases and Materials in International Law* (5th edn, Oxford University Press 2011) 158.

⁴ Hans Kelsen, 14 RCADI 1932/IV 309-310 (Recueil des Cours de l’Académie de Droit International).

⁵ Thomas Grant, ‘Defining Statehood: The Montevideo Convention and its Discontents’ (1998-99) 37 *Columbia Journal of Transnational Law* 404.

⁶ See generally, Grant (n 5); Jean d’Aspremont, *Participants in the International Legal System: Multiple Perspectives on Non-State Actors in International Law* (Routledge 2011).

as a starting point in the determination of statehood⁷ as it provides a succinct and convenient standard to assess whether a community is a state.⁸ Article 1 of the Convention notes that the state as an international person should possess: a permanent population; a defined territory; government; and capacity to enter into relations with other states.

The criterion of having the capacity to enter into foreign relations means independence.⁹ Independence in its plain meaning evinces the ability of a state to be able to dictate and determine its choices and not be subject to the will of another state. This translates into effective control –only then can a state have the capacity to engage in relations with other states. However, this creates difficulties. A country may be legally independent but factually not so long as another state contests its sovereignty. Taiwan, having already met the criteria in Article 1 of the Convention, is legally independent. However, since her independence and sovereignty are still disputed, Taiwan’s status remains ambiguous. This criterion is also deemed problematic because the Convention construes it as a constitutive element of statehood but it can be argued that it is in fact a consequence of statehood. If a state withholds recognition, the entity in question will be deprived of diplomatic relations with that state. An excellent example is the case of Israel in relation to states contesting her statehood. Here we can see again that the Convention is not a straitjacket in the sense that it leaves states considerable discretion when choosing whether or not to recognise states, which downplays the need for an entity to meet a legal criterion to become a state.

Another legal criterion is to have an effective government. It is impossible for any entities to communicate with other

⁷ James Crawford, *Brownlie's Principles of Public International Law* (Oxford University Press 2012) 111.

⁸ Grant (n 5) 414.

⁹ DapoAkande, 'The Importance of Legal Criteria for Statehood: A Response to Jure Vidmar' (10 August 2013) <<http://www.ejiltalk.org/the-importance-of-legal-criteria-for-statehood-a-response-to-jure-vidmar/>> accessed 28 February 2014.

states if there is no functioning government. Rather than being ‘effective’, it is a ‘functioning’ government that allows an entity to become a state – such is the nature of a state. History and current developments have shown that ineffective governments can be characteristic of states. Attempts by states to attract recognition by providing economic incentives (e.g., training of doctors and foreign direct investment) shows that statehood is not a depoliticised process. It appears that despite the efforts to construct a non-political definition of statehood, public international law is unable to provide a purely legal determination. Hence it would be a fallacy to exclude facts (political or historical) that contribute to statehood.

IV. Other possible criteria

Recognition implies an undertaking by the recognising state that it will treat the entity in question as a state.¹⁰ This essay will now move beyond the criteria of statehood articulated in the Convention to examine what motivates a state to recognise another entity and in the process of which, demonstrate that there are indeed both legal and factual criteria for recognition. The criteria to be analysed are i) illegality; ii) absence of competing claims; iii) self-determination; and iv) economic and democratic credentials.

A. Illegality

There are situations where entities seek to achieve statehood by means of an illegal act – a process that is deemed to be in contravention of international law. The criterion of illegality is a barrier to recognition as shown by the international community’s ethical unanimous rejection of Turkish Northern Cyprus’ claim to statehood. This example shows that only law, and not fact, creates a state. However, it is argued that facts and law can be both relevant in illegality. In this regard, there is the sterling example of the North Atlantic

¹⁰ Crawford (n 7) 144.

Treaty Organisation's (NATO) bombing of Yugoslavia, which was done despite the absence of atrocities and massacres at the point of bombing. Despite having the genesis of a violation of Article 2(4) of the UN Charter, Kosovo is not stumbled from doing what she seeks to achieve today – statehood. It is political and historical facts like these that allow Kosovo to emerge as a separate entity from Yugoslavia and begin to court recognition as a state, a legal process.

B. Absence of competing claims

According to the *Restatement (Third)*, an entity must claim to be a state for it to be a state.¹¹ According to state practice in secession, the legal status of this particular entity will remain ambiguous until the competing claims are relinquished.¹² This is clearly a fusion of both factual developments and legal regimes working together to cement sovereignty of an entity. Geopolitical facts can be the decisive factor for recognition. Unlike Taiwan, Bosnia received recognition readily, which was ensured by the removal of the Socialist Federal Republic of Yugoslavia (SFRY)'s claim to territorial integrity¹³ while East Timor was created with Indonesia's consent.¹⁴ The geopolitical realities facing Taiwan are radically different from Bosnia and East Timor. Despite meeting the criteria in Article 1,¹⁵ the translation for the Economic Cooperative Framework Agreement inked between China and Taiwan in 2010 (两岸经济协议),¹⁶ which translates to Economic

¹¹ American Law Institute, *Restatement of the Law (Third): the Foreign Relations Law of the United States*, 201.

¹² Jure Vidmar, 'Explaining the Legal Effects of Recognition' 61 *International and Comparative Law Quarterly* 361.

¹³ Jure Vidmar, 'The Importance of Legal Criteria For Statehood – A Rejoinder to Dapo Akande' (EJIL: Talk!, 9 August 2013) <<http://www.ejiltalk.org/the-importance-of-legal-criteria-for-statehood-a-rejoinder-to-dapo-akande/>> accessed 28 February 2014.

¹⁴ Jure Vidmar, 'Territorial Integrity and the Law of Statehood' (2013) 44 *George Washington International Law Review* 135.

¹⁵ Additional criteria are democracy and economic value to other states.

¹⁶ Liǎng'àn Jīngjì Xiéyì.

Agreement Between Two Shores)¹⁷ illustrates how competing claims remains an impediment to Taiwan's international recognition despite compliance with the Convention. The removal of this impediment is necessary for a legal process of recognition to be forthcoming.

C. Self-determination

Claims to self-determination centre upon the hopes of peoples within a particular entity for independence, which is usually fuelled by nationalist fervour. State practice has shown that recognition has been given on such ground and was recently held not to be in contravention to international law.¹⁸ Indeed, some scholars claim that denial of self-determination is contrary to *juscogens*. However, self-determination's usefulness as a legal regime during the decolonisation era, where colonies could rely on it in order to pursue statehood, is unquestionable. Today, where the geopolitics have changed, self-determination has, to a large extent, lost its *raison d'être*. Hence this is a legal regime that will not enable a state to unequivocally rely on in order to pursue statehood.

The claim to self-determination has been relied upon by East Timor successfully,¹⁹ and it was invoked by Kosovo a few years ago. In the Advisory Opinion of the International Court of Justice, the question to be considered was whether the declaration of independence was in accordance with international law. In order to answer this question, the court began by examining the relatively recent history of Kosovo,²⁰ to eventually conclude that the act did not contravene

¹⁷ Choice of "two shores" rather than states or names of the states reveal no recognition is given.

¹⁸ Advisory Opinion by the International Court of Justice, 22 July 2010: Accordance with international law of the unilateral declaration of independence in respect of Kosovo.

¹⁹ Separate Opinion of Judge Oda, *East Timor (Portugal v Australia)* ICJRep 1995, 90.

²⁰ Accordance with international law of the unilateral declaration of independence in respect of Kosovo (Request for Advisory Opinion), 22 July 2010, 57.

international law. A reason why the declaration of independence did not contravene international law was because Kosovo had self-governing elements in place which were enabled by the UNMIK regulation 2001/9 of 15 May 2001 on a Constitutional Framework for Provisional Self-Government. It is notable that self-determination did not make practical headway in Kosovo's case, which is for Serbia to relinquish her claim on Kosovo. However the main takeaway is that the declaration of independence is only necessary to the extent required to formalise statehood. For this reason, a legal regime alone will not be sufficient for a state to achieve recognition and additional factual credentials serving as incentives will be necessary for recognition.

D. Economic and democratic credentials

For an argument that formation of a state is largely factual, *effectivité* is deemed to be the dominant principle in gaining recognition.²¹ Enhancing economic and democratic credentials are good strategies that states can adopt towards winning recognition thus bettering a state's external *effectivité*.²² A reason why an entity has not received recognition could be that an impetus has not become apparent to the recognising state. This reason could come in the form of economic incentive. This is especially so if the two states have nothing in common –(e.g., cultural, geographical proximity or historical background). States that recognise the sovereignty of Taiwan are Pacific island-states and Caribbean island-states that have no immediate impetus to establish diplomatic links but have done so after fiscal negotiations and developmental training poured in. At the moment, the number of states recognising Taiwan is inversely proportionate to the economic and political clout of China and this is expected to widen. This reveals how

²¹ James Crawford, 'The Criteria for Statehood in International Law' (1976) 48(1) British Yearbook of International Law 93, 95.

²² Professor d'Aspremont mentioned this when musing on how Kosovo could better its external *effectivité*: d'Aspremont and Lielländer (n 15) 18.

sensitively states can respond to incentives. Unsurprisingly, Kosovo has followed suit.²³ With practical benefits, states can readily find a reason to extend recognition. Hence an entity that has met the Convention's criteria might have to rely on additional criteria to factually achieve what the Convention can legally affirm.

Grant noted the accretion of state practice indicating the criterion of democracy as a pre-requisite to statehood.²⁴ Crawford noted that states hesitated to recognise Guinea-Bissau after its unilateral declaration of independence from Portugal, which was, in part, motivated by concerns over the undemocratic character of the new regime. Following the break-up of the Soviet Union, democracy is unquestionably attached with greater importance. To recognise secessionist republics, the EU and the US required that the aspirants to statehood undertake democratic reform. Hence, it is generally the case that a government that is deemed democratic by the international community will enjoy greater international *effectivité*. Likewise, a democratically elected government typically enjoys stronger internal *effectivité*.²⁵ However this will probably only hold true at the point or period when entities are seeking recognition as it is observable that illiberal democracies exist. Hence, a state need not be a true democracy in the ideal sense after it receives recognition, and it is also unlikely that its recognition as a state will be diminished.

A *quid pro quo* thus exists at state level on the international plane. Even if it is not a *quid pro quo*, there has

²³ Foreign Affairs Minister Hoxaj visited Singapore in January 2013 to make a request for recognition. He met the Minister for Law and Foreign Affairs, and discussed possible future collaborations on "diplomacy, education and urbanization" Ministry of Foreign Affairs, 'Minister Hoxhaj in Singapore, seeks recognition of Kosovo' (29 January 2013) <<http://www.mfa-ks.net/?page=2,4,1573>> accessed 1 November 2013.

²⁴ Thomas Grant, *The Recognition of States: Law and Practice in Debate and Evolution* (Praeger 2007)94.

²⁵ Jean d'Aspremont, "Post-Conflict Administrations as Democracy-Building Instruments", (2008) Vol 9 Chicago Journal of International Law", 15.

to at least be an impetus or qualifying reason (democracy) for a state to recognise a particular entity. From the point of seeking recognition to the result of attaining recognition, it is actually a process that begins politically and progresses to a legal act, ending as a legal fact. Hence constitutive and declaratory theories must be accepted as representative of state practice.

V. Concluding remarks

Recognition is motivated by facts that provide incentives or an impetus. In the same vein, historical and political facts could be stumbling blocks for recognition. In this vein, the law is but a means to cement the sovereignty of a state. While non-recognition and the absence or lack of recognition may be motivated by a refusal to accept the legality of another entity, it may also be a result of inaction due to the absence of an impetus that may come in the form of incentives; consequently, this speaks of the inconclusive nature of the legal regimes that purport to qualify statehood. If political facts render recognition to be necessary, convenient or purposeful, then there is no need for legal regimes. However the analysis does not stop here. Even if recognition begins as a factual process, what follows to cement recognition are legal acts that become legal facts. Hence the process of recognition is a progressive one that begins usually with facts favouring recognition, which then usher in the law that cements recognition. State practice also revealed that recognition is a highly discretionary act: therecognising state will decide whether to and when to recognise an entity as a state, therefore underscoring how crucial the facts that provide incentives or an impetus to bequeath recognition are.

Whether there should be less discretion or flexibility is a non-issue because if states are sovereign, there should not be a curb on their discretion although this can worryingly mean ignoring (or approving) illegality. International law will always have a political element to it and the legal regimes of statehood will not be averse to it as well. What makes an entity a state will vary according to the era, geography,

history, and mutual benefits that the recognising state and the entity have between them, factors that all tie in with political facts. As a result, recognition, which is constitutive in nature, is also declaratory in nature.

Best Interests cannot be divorced from Family Faith

Stephanie Murphy

Abstract

English courts are often required to grapple with the dilemma of medical decision-making on behalf of a young child in the situation where conflict arises between parental opinions and between the opinions of parents and healthcare professionals. The approach to these decisions focuses on the immediate best interests of the child as medical priorities. This essay seeks to argue that this approach is short sighted and narrow in the context of parental religious objections to treatment. The current approach can have wider long-term effects on families, and case studies of Re A (Conjoined Twins) and parental refusals from Jehovah's Witnesses will be used to illustrate this. In light of religious objections to treatment seemingly being dismissed by judges all too easily, this essay argues that a wider view of best interests must be embraced.

I. Introduction

This essay will demonstrate how the approach to medical decision-making on behalf of young children in England and Wales focuses on the immediate best interests of an individual.¹ It is argued that this approach is short sighted and narrow, particularly when parental religious views come into conflict with the child's best interests. The application of the best interests test, notwithstanding religious objections from parents, will therefore be the focus of this essay. Case studies of *Re A (Conjoined Twins)*² and parental refusals from Jehovah's Witnesses will be used as illustrations. It is submitted that the long-term wider effects on families as a whole should be awarded greater consideration. In order for the law to achieve this, it requires recognition that the best

¹ Due to the parameters of this essay, competent children, or those of sixteen and seventeen years of ages, will not be discussed.

² [2001] Fam 147.

interests of a child are more than just medical interests. A wider view of best interests must be embraced in the context of religion, as it appears that some judges too easily dismiss objections stemming from this context.

II. Best Interests as Medical Priorities

In the overwhelming majority of situations, parents have a decisive role in medical decision-making on behalf of their children who are unable to express any views of their own. However, this is not always the case. The courts are sometimes required to grapple with the dilemma of medical decision-making when conflict arises between the parents themselves and the views of parental and healthcare professionals. It is evident from the existing case law that in these types of cases treatment decisions turn on an assessment of the child's immediate best interests. The immediate best interests appear to constitute what is best in terms of the child's medical condition.

The benchmark of determining best interests was first considered in *Re B*.³ This case concerned the court overriding parental refusal to consent to an operation that would remove an intestinal blockage from their child who suffered from Down's syndrome. The court ultimately rejected the argument that the views of responsible and caring parents must be respected. The decision to operate was made in the child's best interests, contrary to the parental views. Approvals may be sought from the Family Division of the High Court if parents refuse to provide consent.⁴ The best interests approach was confirmed in *Re J (A Minor)*⁵ where it was stated by Lord Donaldson that "there is a balancing exercise to be performed in assessing the course to be adopted in the best interests of the child".⁶

³ [1981] 1 WLR 1421.

⁴ Children Act 1989, s8.

⁵ [1990] 3 All ER 930.

⁶ *ibid* 938.

This approach continues to be adopted, as is evident from recent case law.⁷ It is clear, therefore, that there has been broad judicial endorsement of the best interests approach. This suggests that it is held as an appropriate standard and it still applies notwithstanding religious objections from a child's parents.⁸ However, whether this should be the case is questionable.

The characteristics of the case law demonstrate, through the current approach to the test, that the child's immediate health is prioritised. It can be inferred that a best interests approach is the law's attempt to protect vulnerable individuals - namely, young children who lack the ability to make decisions for themselves. By focusing on immediate best interests, the child's medical condition is put at the forefront of the decision. Healthcare professional opinions are usually adhered to, and as Kearney notes, "[m]any doctors tend only to be guided by medical priorities."⁹

This approach can be viewed as consistent with professional guidelines and statutory authority: the GMC's requirement that the "primary duty is to the child",¹⁰ and the Children Act which states that "the child's welfare shall be the court's paramount consideration".¹¹ When examining the courts' judgements, the outcomes have been achieving these aims. For example, in the above stated cases the medical condition of the child has always been prioritised. However, arguments that this approach is an appropriate standard are not always convincing and there are some inconsistencies.

III. Inconsistency in the Law

There is inconsistency in the law and the way in which the best interests test is applied which may suggest that some

⁷ *An NHS Trust v SR* [2012] EWHC 3842 (Fam).

⁸ *Re R* [1993] 2 FLR 757.

⁹ PJ Kearney, 'Leukaemia in Children of Jehovah's Witnesses: Issues and Priorities in a Conflict of Care' (1978) 4(1) *Journal of Medical Ethics* 34.

¹⁰ GMC, *Treatment and Care Towards the End of Life: Good Practice in Decision Making* (2010) [90].

¹¹ Children Act 1989, s1(1).

judges are more wary of religious than other objections. An illustration of this can be seen in *Re T (A Minor)*,¹² which concerned parental refusals to their child's life-saving liver transplant. They did not perceive the transplant as in their child's best interests. The parents, both healthcare professionals, fled the country in fear of intervention. The court ruled that the parental views should be accepted, despite these views conflicting with the medical priorities of the child. An explanation from the seemingly contrasting decisions concerning the endorsement of parental views was highlighted in the judgement of Waite LJ. He makes a clear distinction between cases where parental opposition is "prompted by scruple or dogma of a kind which is patently irreconcilable with principles of health and welfare widely accepted by the generality of mankind" and "highly problematic cases where there is a genuine scope for a difference of view".¹³

The inconsistency with this case and others is clear. Fox and McHale question this by asking, "...whether the...judges were right to be more receptive to the parents' objections in *Re T*"?¹⁴ Other possible explanations are that more weight was given to parental views since they were healthcare professionals. An additional explanation could be due to the complication of the parents having fled the country. Whatever the reasoning, the decision in this case is unusual. However, the case should not be dismissed as a one-off since it raises broader issues concerning the adequacy of our current framework.¹⁵ Inconsistencies of this case in comparison to those that concern religious objections will be discussed below. The unusual outcome in this case, where parental views have been respected despite being contrary to medical interests, can be used to emphasise the narrowness of the way the best interests standard is applied.

¹² [1997] 1 WLR 242.

¹³ *ibid* 254.

¹⁴ Marie Fox and Jean McHale, 'In Whose Best Interests?' (1997) 60(2) *Modern Law Review* 703.

¹⁵ *ibid* 709.

It has been shown that the policy underlying the law in its current guise, decision-making on the basis of immediate best interests, aims to protect the vulnerable. However, it will now be established that, as the law strives to meet this aim, the best interests approach can be perceived as short sighted and narrow, particularly when faced with religious objections from parents.

It is suggested that the law should protect the vulnerable by looking further into the family unit. Case studies of *Re A (Conjoined Twins)* and the objections to blood transfusions from Jehovah's Witnesses will be used to illustrate that dismissing religious objections from parents, on the basis of a child's best interests, can have wider long-term effects on families as a whole.¹⁶ It is argued not that religious views should always be imposed upon the child, but that a broader approach to medical decision-making should be adopted.

IV. *Re A (Conjoined Twins)*¹⁷

The case of *Re A* concerned a set of conjoined twins whose parents were devout Roman Catholics. Death was inevitable for the twins if they were not separated by surgery. Separation would give the stronger twin a very good chance of survival but would result in the death of the weaker twin. It would have been against the parent's religion to consent to this surgery. Despite parental refusal, the separation was permitted after balancing the interests of the twins and a defence under the doctrine of necessity was utilised.

The best interests approach appears to be under pressure here. Freeman states how "[t]he Court of Appeal...concluded that the operation could not be justified as in the best

¹⁶ Due to the parameters of this essay only these two case studies will be analysed. However, there are other aspects of religious objections from parents that could have equally been examined. For example, refusals to the removal of ventilation (*Re C* [1998] 1 FLR 384) or disputes about religious circumcision (*Re J* [2000] 1 FCR 307) could also have been used to illustrate the narrowness of the best interests approach.

¹⁷ *Re A (Conjoined Twins)* (n 2).

interests of...the weaker twin. Since it was not in her best interests to die".¹⁸

It is obvious that the best interests of each twin were diametrically opposite. Hence, this suggests that taking this approach may have been inappropriate. It is narrow since the best medical interests of the stronger twin were prioritised. Wider issues should have been taken into consideration. For example, the emotional distress on the parents could have been awarded more respect by addressing values they regarded as highly important. As Hewson states, the court "...implicitly downgraded the doctrine of the sanctity of life."¹⁹

Re A emphasises that the court's view of decisions made on the basis of religious conviction is narrow, since the strong parental views were dismissed. Freeman asks the question, "did too much hinge upon the religion and culture of the parents?"²⁰ Gillon convincingly argues that in this case the court "should have allowed the parents to refuse medical intervention", since the parents' moral reasoning was "not eccentric or merely religious, but widely acceptable".²¹ This is convincing because if both reasoning for and against the separation was morally adequate, why should the court's view have been for one to prevail? Parental views should have been accorded more respect.

Re A is difficult to reconcile with the case of *Re T*, which was discussed above. Fox and McHale observe "Re T focussed on broader circumstances."²² Possible reasons for the disparity between this case and others have already been considered. *Re T* is surprisingly inconsistent with religious refusal cases because wider long-term effects were taken into consideration despite "no potential rejection of the child

¹⁸ Michael Freeman, 'Whose Life is it Anyway?' (2001) 9(3) Medical Law Review 276.

¹⁹ Barbara Hewson, 'Killing off Mary: Was the Court of Appeal Right?' (2001) 9(3) Medical Law Review 298.

²⁰ Freeman (n18) 275.

²¹ Raanan Gillon, 'Imposed Separation of Conjoined Twins - Moral Hubris by the English Courts?' (2001) 27(1) Journal of Medical Ethics 4.

²² Fox and McHale (n14) 703.

prompted by religious doctrine.”²³ What is important to draw from *Re T* is that it demonstrates that the best interests approach can be applied in a broader manner than what is best in terms of the child’s medical health. If a broader approach were to be adopted, the inconsistencies that can undermine the law’s coherency would appear less dramatic and more justified.

Michalowski’s argument substantiates the recommendation that a broader approach should be embraced. She emphasises how the twins’ medical condition was prioritised:²⁴

“*Re A* demonstrates that the courts’ approach to substitute their own decision for that of the parents combined with the tendency to give precedence to medical evidence leads to a shift from parental powers to a more and more influential role for the medical profession...physicians cannot be the best judges of their child patients’ best interests.”

It is suggested that more weight should have been awarded to parental views in this case. Furthermore, physicians may not be the best judges of best interests since parents are the ones who are responsible for their children after treatment, know them best and care the most. This can similarly be argued in relation to parental refusals from Jehovah’s Witnesses. However, the reasoning for refusals in these cases is much more controversial.

V. Parental Refusals From Jehovah’s Witnesses

The case study of parental refusals from Jehovah’s Witnesses to consent to treatment involving the use of blood products on their children provides an additional illustration of how

²³ Caroline Bridge, ‘Religion, Culture and Conviction - The Medical Treatment of Young Children’(1999) 11(1)Child and Family Law Quarterly 10.

²⁴ Sabine Michalowski, ‘Reversal of Fortune - Re A (Conjoined Twins) and Beyond: Who Should Make Treatment Decisions on Behalf of Young Children?’[2001] Health Law Journal 168.

the best interests test can be recognised as short sighted and narrow. The reasons for such refusals are very controversial. Based on biblical texts,²⁵ the perceived consequences for Jehovah's Witnesses receiving blood products are catastrophic. It is not suggested that these controversial religious beliefs should always be imposed upon the child, but that decisions should be made having examined potential, wider long-term effects.

There is a line of case law that demonstrates that the courts will override parental rejections of life saving blood products on the basis of a child's best interests. For example, in *Re R*,²⁶ Booth J stated how the court was "bound to override the parents' wishes."²⁷ This was similarly done in *Re S*.²⁸ In *Re S* parental religious views were awarded some recognition since there was a delay in administering treatment. However, as Bridge insightfully points out, "[w]hy...wait so long when they knew...the parents would never agree?"²⁹ This is not an effective method of recognition as the child stands only to suffer and the approach is still narrow.

It is argued that the best interests approach is narrow because decisions are only based on immediate medical interests, without sufficient recognition being awarded to religious beliefs. McHale and Fox question "whether the down playing of such beliefs is legitimate."³⁰ The discretion that parents possess, or at least ought to possess, in how to bring up their child also appears to be ignored. Bridge states that it is a natural instinct of parents to impose "their own set

²⁵ Genesis 9:4, Leviticus 17:10 and Acts 15:29.

²⁶ *Re R (Minor)(Blood Transfusion)*[1993] 2 FLR 757.

²⁷ *ibid.*

²⁸ [1993] 1 FLR 376 Explanatory Notes: S, aged four and a half, suffered from T-cell leukaemia. There was strong medical opinion that in order to maximise the prospect of successful treatment and saving the life of S, the transfusion of blood products was essential. S's parents objected to this treatment due to religious belief and concerns over the safety of the of blood products. Thorpe J overruled the parental objections on the basis of promoting the child's welfare.

²⁹ Bridge (n23) 7.

³⁰ Fox and McHale (n14) 703.

of values in relation to upbringing. It is arguably their right to do so.³¹ Families could present this reasoning to argue that their Convention Rights demand a wider outlook on best interests.³²

A wider outlook requires long-term effects to be considered in addition to immediate medical interests. Kearney states how “decisions have to be considered both in the short and the long term.”³³ He continues to argue that “[d]ecisions are too often based on short-term medical contingencies. This may be counter-productive if the short-term gains are less important than the long-term problems.”³⁴

Kearney is convincing in his argument since he suggests that the medical condition of the child remains vitally important, but more consideration should be awarded to long-term effects such as the disruption of the family unit. Since “destruction of these relationships...often has a detrimental effect on the patient”³⁵ a wider outlook than on immediate best interests is required. It has been reported that a Jehovah’s Witness man, from Ghana, disowned his five-year-old son as a result of him having received a life-saving blood transfusion.³⁶ Although this is an extreme example, it provides an illustration of possible wider long-term effects on families- something that the law in its current guise fails to accommodate.

Wider long-term effects on families appear to fall outside the ambit of immediate best interests. This is likely to have detrimental effects on patients. As Gilbar argues, “[t]he close link between family support and involvement in decision-

³¹ Bridge (n23) 3.

³² Article 8 and 9, European Convention on Human Rights.

³³ Kearney (n9) 32.

³⁴ *ibid* 34.

³⁵ *ibid* 32.

³⁶ ‘Jehovah Witness couple disown five-year-old son for receiving blood’ (Ghana Web) <<http://www.ghanaweb.com/GhanaHomePage/NewsArchive/artikel.php?ID=181322>>accessed on 1 April 2013.

making indicates that family support is one of the main considerations”.³⁷

Although he is not writing in the same context as is being discussed here, his reasoning, that the support of the family is crucial, can be imposed in the situation of parental religious refusals. Adding weight to this argument is Kearney’s statement that “[t]he parents’ fundamental beliefs should not be easily dismissed. A strong family unit is almost essential in the management of children with cancer.”³⁸

It can be inferred that whilst a child is being treated for a serious illness a strong family unit is highly desired. In the context of parental religious refusals, there is a high risk of disruption of the family unit because the perceived consequences of treatment can be grave. Hence, this is an area where greater weight should be awarded to parental views.

The best interests approach has shown to be narrow and short sighted in relation to religious objections from parents, as the approach appears to have a sole focus on medical priorities. It is stated by Bridge that “[a] case can persuasively be made out for upholding the parental religious interests for, on balance, the child’s welfare in the context of his culture, religion and family is enhanced”.³⁹

Therefore, although a child’s medical welfare will be central, other interests, such as the impact on family relations, should be taken into consideration. This is important because wider long-term effects can sometimes be more detrimental than not acting in, what the law may currently perceive to be, the immediate best interests.

VI Conclusion

It is evident from the existing case law that the approach taken in medical decision-making on behalf of young

³⁷ Roy Gilbar, ‘Family Involvement, Independence and Patient Autonomy in Practice’(2011) 19(2) Medical Law Review 19(2) 225.

³⁸ Kearney (n9) 34.

³⁹ Bridge (n23) 5.

children in England and Wales focuses on the immediate best interests of the child. The purpose of this is to protect the vulnerable. When decisions are made in the best interests of a child, which conflict with strong religious views held by the parents, this approach has proven to be short sighted and narrow. The adoption of an approach that encapsulates and recognises more than just medical interests is recommended. This is more appropriate as greater consideration would be awarded to the wider long-term effects of families who have religious objections. A wider view of best interests must be embraced.

Different categories of employment status: an opportunity for employers to cut costs or another hurdle to overcome?

Sophie Cooke

Abstract

There are different categories of employment status found in the workplace: employees, workers and independent contractors. The employment status of an individual determines which employment rights they will benefit from, and in turn, which obligations their employer will be subject to. The availability of these different categories allows employers financial flexibility and can be used to their advantage to cut costs. This article examines how the intervention of the courts in determining employment status has made it more difficult for employers to make commercial choices when employing individuals in order to keep costs down.

The tests used to determine employment status are examined, highlighting the characteristics typically found in each category. I then discuss the obligations that each category imposes on employers and how this might affect their costs. Through an examination of the courts' intervention in determining employment status, particularly in relation to vulnerable workers, it will be argued that they have made this area of law too uncertain, and how, as a result, any potential savings that employers could derive from these categories is sacrificed.

I. Introduction

Traditionally two categories of employment status have been recognised: employees and independent contractors. While judges have struggled for years to classify labour relations, this so-called binary divide is a development of the 20th century.¹ An employee is “an individual who has entered into or works under a contract of employment”² whilst an independent

¹ Simon Deakin, ‘Does the “Personal Employment Contract” provide a basis for the reunification of employment law?’ (2007)36(1) *Industrial Law Journal* 68.

² Employment Rights Act 1996, s 230(1).

contractor is self-employed, working under a contract for services.

Evidently these definitions are not particularly prescriptive. As a result, the courts have developed various tests to determine employment status by examining aspects of the relationship. Where the employer controls, or has the right to control, how the work is done, it is likely that he is dealing with an employee.³ The greater the degree of control, the more likely the court is to find a contract of service.⁴ By the organisational test, an employee is well-integrated and considered “part and parcel of the organisation”.⁵ Mutuality of obligations is considered an essential pre-requisite for an employee. This is satisfied where there is an obligation on the employer to provide work and on the individual to accept it.⁶ Alternatively, if the economic reality of the situation suggests that individuals are in business on their own account, bearing the financial risk and enjoying the profits, then this usually makes them an independent contractor.⁷

In light of the wide range of factors that should be taken into consideration when determining employment status, the courts have more recently developed a multi-factorial approach, whereby a variety of factors are considered together. At a minimum, control and mutuality of obligations must be present for an individual to be classed as an employee.⁸

Since 1997 an intermediate category has been recognised: the ‘worker’. This is:⁹

³ *Yewen v Noakes*[1880] 6 QBD 530; *Walker v Crystal Palace Football Club*[1910] 1 KB 87.

⁴ *Whittaker v Minister of Pensions and National Insurance*[1967] 1 QB 156.

⁵ *Bank voor Handel en Scheepvaart NV v Slatford*[1953] 1 QB 248, 295.

⁶ *Montgomery v Johnson Underwood*[2001] EWCA 318 (Civ).

⁷ *Market Investigations Ltd v Minister for Social Security*[1969] 2 QB 173.

⁸ *Ready Mixed Concrete (South East) v Minister for Pensions and National Insurance*[1968] 2 QB 197.

⁹ *Employment Rights Act 1996*, s 230(3).

“An individual who has entered into or works under a contract of employment, or any other contract...whereby the individual undertakes to do or personally perform any work or services for another party to the contract whose status is not by virtue of the contract that of a client or customer of any profession or business undertaking carried on by the individual.”

By this definition, all employees are workers. However, worker is a wider concept; the category was introduced to extend protection to people who are excluded from the narrow definition of an employee, but are not in business on their own account and therefore require protection due to their economic dependency on their employer.¹⁰ Again the courts have identified a number of key factors within this definition. The individual must contract to perform the work personally, there must be mutuality of obligations and the individual must not be in business on their own account.¹¹ This additional third category catches individuals who do not satisfy the definition of an employee, but as they are not in business on their own account, are not truly independent contractors. A further distinction will also be made between typical and atypical workers. The label refers to their mode of work and whether the working arrangement is considered ‘typical’. It is important to note that both typical and atypical workers can still fall within any of the three categories of employment status.

The law distinguishes between these different categories within the workplace in order to regulate the application of protective legislation, common law rights and for tax and National Insurance purposes. This is important for employers because it is the different rights and responsibilities that accrue to each group that affects their business costs.

¹⁰ Guy Davidov, ‘Who is a worker?’ (2005) *Industrial Law Journal* 57, 57.

¹¹ *Byrne Bros (Formwork) Ltd v Baird and Others* [2002] ICR 667.

II. How different categories of employment status provides employers with opportunities to reduce costs

A. *Variations in statutory obligations*

Statute provides for a number of rights to protect individuals in the workplace and compliance with these rights can be expensive for employers. These rights do not apply equally across the different categories, as the Employment Appeals Tribunal explains:¹²

“The reason why employees are thought to need such protection is that they are in a subordinate and dependent position vis-à-vis their employers: the purpose of the Regulations is to extend protection to workers who are, substantively and economically, in the same position. Thus the essence of the intended distinction must be between, on the one hand, workers whose degree of dependence is essentially the same as that of employees and, on the other, contractors who have a sufficiently arm's-length and independent position to be treated as being able to look after themselves in the relevant respects.”

All three categories are protected under equality legislation.¹³ This is the only statutory protection accorded to independent contractors, however it should not be overlooked as failure to properly comply can result in costly litigation; a successful claim for equal pay can require an employer to retrospectively equalise the pay for up to 6 years.¹⁴ Workers benefit from some limited additional protection, which includes entitlement to the National Minimum Wage,¹⁵ protection under the Working Time Regulations¹⁶ and entitlement to Trade Union

¹² *Byrne Bros* (n 11) 677.

¹³ Equality Act 2010.

¹⁴ *ibid* s 132.

¹⁵ National Minimum Wage Act 1998, s 1.

¹⁶ Working Time Regulations 1998, SI 1998/1833, reg 2.

membership.¹⁷ Employees benefit from all the rights provided for by statute, including family and unfair dismissal rights that the other categories do not enjoy.¹⁸

Employers inevitably incur costs in complying with these rights. Trade Unions “often secure better terms and conditions of employment”¹⁹ through collective bargaining, such as improved social rights and higher pay.²⁰ They can also organise industrial action²¹ which is disruptive for an employer who will see a decline in productivity if a large portion of his staff are on strike. Minimum wage increases annually,²² irrespective of the economic conditions that employers may be facing. Although the increase is often only a matter of pence, across a whole business this can be costly for employers. Furthermore, while employers can claim back at least 92% of maternity pay, which may rise to 100% depending on National Insurance contributions,²³ there is no automatic entitlement relating to sick pay where employers are left out of pocket.²⁴ They may also face additional costs in hiring and training a temporary replacement.

Indeed an argument often cited by those against regulation is that the cost of compliance is high. Employers have stated that “employment law does have a negative effect on business growth, and the concern... is about cost, complexity and the ... cost of ‘getting it wrong’”.²⁵ For this

¹⁷ Trade Union and Labour Regulations (Consolidation) Act 1992, ss 1, 29.5 and 29.6.

¹⁸ Employment Rights Act 1996, ss 71 and 94 for maternity leave and unfair dismissal respectively; Social Security Contributions and Benefits Act 1992, s 151 for statutory sick pay.

¹⁹ Hugh Collins, KD Ewing and Aileen McColgan, *Labour Law: Text and Materials* (2nd edn, Hart Publishing 2005) 723.

²⁰ Collins (n 19) 647.

²¹ Trade Union and Labour Regulations (Consolidation) Act 1992, pt V.

²² Government Digital Service, ‘National Minimum Wage rates’ (GOV.UK) <<https://www.gov.uk/national-minimum-wage-rates>> accessed 16 November 2013.

²³ HMRC, ‘Employer Helpbook for Statutory Maternity Pay’ (2013) E15, 30.

²⁴ HMRC, ‘Employer Helpbook for Statutory Sick Pay’ (2013) E14, 31.

²⁵ The Better Regulation Executive, ‘Lightening the Load: The Regulatory Impact on UK’s Smallest Businesses’ (November 2010)

reason, it is beneficial that an employer can, where appropriate for the role, hire workers or independent contractors rather than employees, circumventing increased costs that may have arisen from compliance with statutory obligations or collective bargaining.

B. Costs relating to implied duties

In addition to statutory protection, employees benefit from a number of implied duties to which the employer is subject. These duties do not apply to workers or independent contractors. Employees may enforce these even where employers do not intentionally contract to them. The implied duty of good faith²⁶ prohibits conduct likely to produce destructive or damaging consequences.²⁷ A range of conduct can result in breach of this duty, such as use of foul language²⁸ or failure to provide a grievance procedure,²⁹ making the employer liable to pay damages. This additional financial burden in respect of employees strengthens the case for using workers and independent contractors as attractive cost-saving alternatives.

C. Costs arising from vicarious liability

Another crucial distinction affecting employers' costs is found in the doctrine of vicarious liability, by which an employer is liable for the torts of his employee:³⁰

“Every act done by a servant in the course of his duty is regarded as done by his master's orders and consequently is the same as if it were the master's own act.”

<https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/31614/10-1251-lightening-the-load-regulatory-impact-smallest-businesses.pdf> accessed 16 November 2013.

²⁶ Imperial Group Pension Trust Ltd v Imperial Tobacco Ltd [1991] 1 WLR 589.

²⁷ University of Nottingham v Eyett and Another [1999] ICR 721.

²⁸ OgilvieNeyrfor-Weir Ltd (EAT, 15 May 2003).

²⁹ WA Goold (Pearmak) Ltd v McConnell [1995] IRLR 516.

³⁰ Bartonshill Coal Co v McGuire [1858] 3 Macq 300.

The court has given the determinative concept of ‘course of employment’ a broad interpretation to cover acts closely connected to the employee’s work.³¹ Consequently employers can be liable for actions that they have no control over and which often occur outside working hours and the workplace.³² This includes situations where the act has been expressly forbidden,³³ was an intentional tort³⁴ or even criminal activity.³⁵

Recent case law has given rise to a new precedent, which suggests that workers may also benefit from the doctrine of vicarious liability. Traditionally the employer would have to exercise sufficient control over the individual in order for their relationship to give rise to vicarious liability.³⁶ This has recently been expanded to include relationships that are “akin to that of employer and employee”.³⁷ As the tests for a worker replicate some of the tests for an employee, it is possible that the court could find the employer-worker relationship sufficiently analogous to give rise to vicarious liability. By comparison, concerning independent contractors, employers are only liable where they are subject to a non-delegable duty of care,³⁸ arising where the activity involved is particularly hazardous³⁹ or where the employer has assumed responsibility.⁴⁰ Where vicarious liability applies, the employer will have to pay damages to the victim. Indeed, a regularly cited justification for this imposition of strict liability is that the employer has deeper pockets and can better afford the compensation.⁴¹ However, this can amount to substantial sums of money and in order to avoid these

³¹ *Poland v John Parr & Son*[1927] 1 KB 236.

³² *ibid.*

³³ *Plumb v Cobden Flour Mills*[1914] AC 62.

³⁴ *Poland* (n 31).

³⁵ *Lister v Hesley Hall Ltd*[2001] UKHL 22.

³⁶ *Honeywill and Stein Ltd v Larkin Bros Ltd*[1934] 1 KB 191.

³⁷ *Various Claimants v Catholic Child Welfare Society*[2012] UKSC 56.

³⁸ *Cassidy v Minister of Health*[1951] 2 KB 343.

³⁹ *Black v Christchurch Finance Company Ltd*[1894] AC 48.

⁴⁰ John Murphy, *Street on Torts* (12th edn, OUP 2007) 600.

⁴¹ Murphy (n 40).

costs an employer may instead choose to hire an independent contractor, ensuring they will only be liable in limited circumstances. From the above it can be seen that the various employment classifications have vastly different rights and responsibilities with the ability to significantly reduce an employer's costs if utilised correctly.

III. Employers, Beware

The prudent employer will use the employment classification most suited to the role to avoid contracting for additional rights beyond what is needed for the work to be completed. Often employers will expressly define their relationship in a way that will not give rise to excessive costs. However, problems arise when the employment status is disputed.

A. Tests lack clarity/coherence

As previously noted, vague statutory definitions have led to the adoption of multiple tests, which may be used individually or combined together in the multifactorial approach.⁴² However, no guidance has been given on how much weight should be accorded to each factor. As a result, different tests may produce different results on the same set of facts. For example, in *Durcan*,⁴³ the claimant was a self-employed dentist, but also worked rota at a local hospital. When his status regarding the hospital work was in dispute, the court found that he was an employee. Factors supporting this included his use of hospital premises and equipment, lack of choice over who to treat and that his remuneration was not dependent on the number of patients treated. However, several other factors pointed towards an independent contractor status. The dentist paid his own tax and National Insurance contributions, had a limited ability to replace himself, and the hospital lacked control over how his work was performed. Had the court followed any one of these tests on its own, it is likely that it would have also found

⁴² See n8 and accompanying text.

⁴³ City and East London FHS Authority v Durcan [1996] EAT/721/96.

him to be an independent contractor. This illustrates the ambiguity surrounding employment status.

Substitution clauses were previously treated as conclusive of a contract for services,⁴⁴ where the individual has the right to substitute themselves, there is no personal service and therefore no mutuality of obligations. However this created problems of ‘sham contracts’,⁴⁵ where the substitution clause would be very limited and in practice the individual would have no such ability. These are used against vulnerable individuals with no real bargaining power in order to circumvent their employment rights.⁴⁶ As a result the court has held that limited substitution clauses do not negate a contract of employment.⁴⁷

To avoid this uncertainty, employers will write comprehensive contracts that clearly indicate employment status, but over time may cease to reflect the reality of their situations and as a result may be struck down by the court. While beneficial to individuals being exploited by their employer, it is problematic for employers who are often not actively seeking to exploit their staff, but simply trying to stay afloat during difficult economic times. They may believe they are dealing with an independent contractor because that is what they contracted to, only to later find out that they are not, and have become subject to all the responsibilities normally associated with an employee, thereby incurring greater costs.

B. Difficulties with atypical workers

Employment status becomes still more confusing in relation to atypical workers. These individuals may not receive regular permanent work, do not work on premises or have

⁴⁴ Express & Echo Publications Ltd v Tanton [1999] ICR 693.

⁴⁵ For more on sham contracts, see Sam Middlemiss, ‘The legal impact on employers where there is a sham element in contracts with their workers’ (2012) 54(3) International Journal of Law and Management 209.

⁴⁶ Alan Bogg, ‘Sham self-employment in the Supreme Court’ [2012] 41(3) Industrial Law Journal 328.

⁴⁷ *Byrne Bros* (n 11).

no single employer. The label ‘atypical worker’ is misleading as often the individuals lack certain attributes of a worker leaving their status uncertain. In particular, mutuality of obligations is necessary for worker status yet this is often absent for casual workers. Arguably the traditional conception of a nine-to-five worker is no longer typical⁴⁸ and in recent years employers have relied more heavily on atypical workers to save costs throughout the recession, with the use of zero-hour contracts increasing by 174% between 2007 and 2012.⁴⁹ These allow employers to respond to fluctuations in demand; they have staff available when needed but who are not guaranteed work, nor paid permanent salaries.⁵⁰ Practitioners claim that employers in other struggling economies envy this flexibility.⁵¹ However an alternative view of this practice is that shifting “the risks of business fluctuations from the employer on to the worker”⁵² is exploitative, as the lack of guaranteed work creates financial instability for these workers. To prevent this, the court will at times take a protective stance. Even where the contract expressly states the employment status, the court will disregard it if it does not believe that it reflects the reality of the situation “...the true agreement will often have to be gleaned from all the circumstances of the case, of which the written agreement is only a part”.⁵³

⁴⁸ Ian Smith and Aaron Baker, *Smith & Wood’s: Employment Law* (10th edn, OUP 2010) 56.

⁴⁹ IDS Thomson Reuters, ‘Zero hours, but not necessarily zero responsibility’ (16 April 2013) <<https://ids.thomsonreuters.com/employment-law/features-analysis/zero-hours-not-necessarily-zero-responsibility>> accessed 20 November 2013.

⁵⁰ *ibid.*
⁵¹ <<http://www.bbc.co.uk/news/business-23573442>> accessed 20 November 2013.

⁵² Sandra Fredman, ‘Labour Law in Flux: The Changing Composition of the Workforce’ (1997) 26(4) *Industrial Law Journal* 337, 337. This article examines the perspective of employers, but for more on the effect on individuals see BBC News, ‘Q&A: What are zero-hour contracts?’ (5 August 2013) <<http://www.bbc.co.uk/news/business-23573442>> accessed 20 November 2013.

⁵³ *Autoclenz Ltd v Belcher and others* [2011] UKSC 41, [2011] ICR 1157, 1168.

A contract may preclude mutuality of obligations, however it has been found to exist in practice where, over a long period of time, a casual worker actually worked on a regular basis.⁵⁴ Where a group of staff on zero hour contracts provided critical around-the-clock care for a severely disabled patient, they were employees because the work was so challenging and personalised that a series of *ad hoc* contracts with independent contractors would have been inappropriate.⁵⁵ The decision of the court to ignore a contract, freely entered into, is somewhat paternalistic. The individual in question may have chosen that arrangement to suit their own lifestyle, and therefore benefit from it as much as their employer does, until something goes wrong. Ignoring contractual stipulations, particularly in zero hour contracts, is just another way of preventing employers from cutting their costs in a difficult economy.

C. Potential expansions of implied duties to workers

Brodie argues that the courts may further extend protection to workers using implied duties. The emphasis on personal service with regards to workers makes them more similar to employees than to independent contractors. On that basis, the courts may well decide to extend the implied duties of employers to include workers, particularly the duty of mutual trust and confidence, which is only appropriate where a contract is personal in nature.⁵⁶ This would impose a greater burden on employers and could be seen as a step too far on the part of the courts. If parliament had intended that workers should be treated as employees, they would have simply legislated to expand the definition of an employee.

⁵⁴ *Cornwall County Council v Prater* [2006] EWCA Civ 102.

⁵⁵ *Pulse Healthcare Ltd v Carewatch Care Services Ltd & Others* (EAT, 6 August 2012).

⁵⁶ Douglas Brodie, 'Employees, workers and the self-employed' (2005) 34(3) *Industrial Law Journal* 253. For more on the extension of the worker category see Davidov (n 10).

IV. Conclusion

Employees, workers and independent contractors are subject to different rights and responsibilities regarding their employment relationship, which can directly affect the cost to employers. Knowledge of these rights can enable the employer to make the most efficient use of his workforce, such as reducing his wage bill through zero hour contracts, or removing liability through the use of independent contractors. However, the unclear and ambiguous tests for establishing employee status make it practically impossible to accurately categorise these individuals and blur the distinction between each group. As a result, employers are unable to know with any certainty which rights and liabilities they are exposed to and the associated costs they must factor in. This is exacerbated by the increasingly protectionist stance of the courts; whilst protecting individuals from exploitation is a valid concern, employers are being prevented from using these statuses to cut their costs in difficult economic times. Indeed, there is little benefit in the availability of these different categories if the courts are too ready to blur the distinctions between them. Therefore, while these different statuses appear to be a useful economic tool, they can equally impose greater hurdles to overcome, resulting in increased costs to the employer. As mentioned above, the cost of getting it wrong is high; the courts need to provide some clarity on the position on employment status to the benefit of both parties.

The CESL proposal: Progress or misstep?

Sam Wardleworth

Abstract

The state of European consumer protection is incoherent and inconsistent across member states. The Commission believe that this has a detrimental effect on cross border trading and costs the EU €26 billion a year in dissuaded trade. In response it proposes a Common European Sales Law, the aim of which is to inspire confidence in the European market, increase competition and ultimately improve the consumer market for all European consumers. However as it stands the CESL is problematic and falls short of this ambition. This essay considers some of its faults from the point of view of B2C transactions, focusing on the scope, features and cost of the CESL. Ultimately, while the CESL has its benefits, it is too inconsistent to have a positive effect on the Consumer Contract acquis, and the threat to existing harmonisation is simply too great to justify the few advantages it offers in its current form.

I. Introduction

As it stands the Consumer Contract acquis is a rather confusing picture and can “hardly be said to be coherent”.¹ In the Commission’s view, this confusion has a substantial negative effect on cross border trading.² In fact, they believe that acquiring information on and adapting to the contract laws of other member states are two of the most significant hurdles that traders face when attempting to trade within the European Union. Clearly businesses would undoubtedly

¹ Nobert Reich, ‘Harmonisation of European contract law: with special emphasis on consumer law’ (2011) 1 China-EU Law Journal 55, 70-76.

² European Commission, ‘European Contract Law in Consumer Transactions: Analytical Report’ (2011) Flash Eurobarometer 321 - The Gallop Organization <http://ec.europa.eu/public_opinion/flash/fl_321_en.pdf> accessed 4 April 2013.

prefer a “legal environment with less complexity”.³ As a result of this impetus, in October 2011 the Commission released a long awaited proposal for a “pan-European”⁴ contract law, known as the Common European Sales Law (CESL). The CESL establishes a uniform set of contract rules that purport to offer a high level of consumer protection indiscriminately. The Commission believes that it will “stimulate the commerce between Member States”⁵ and combat the €26 billion lost each year to dissuaded trade.⁶ This will improve competition and therefore potentially lead to a greater choice and lower prices for over 500 million European consumers.

This essay focuses on Business-to-Consumer (B2C) transactions and as such we must consider the *real* nature of a consumer. Critically, consumer confidence is aesthetic; a consumer will have more confidence if the market becomes strengthened, through experience and word of mouth.⁷ It is an exceptional consumer who surveys the contractual relationship when making purchases.⁸ That is not to say that the consumer does not stand to benefit from the CESL. If the CESL makes cross-border trading more attractive, then it could have a positive effect on consumers by increasing

³ Nicole Kornet, ‘The Common European Sales Law and the CISG: Complicating or Simplifying the Legal Environment?’ (2012) 19 *Maastricht Journal of European and Comparative Law* 164.

⁴ Pinsent Masons, ‘MPs back Government over Common European Sales Law stance and question legal basis for proposed regime’ (2013) <www.out-law.com/en/articles/2013/january/mps-back-government-over-common-european-sales-law-stance-and-question-legal-basis-for-proposed-regime> accessed 4 April 2013.

⁵ Directorate-General For Internal Policies, ‘Consumer Protection under the Proposal for a Common European Sales Law’ (2012) <[http://www.europarl.europa.eu/RegData/etudes/note/join/2012/462507/1/POLJURI_NT\(2012\)462507_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/note/join/2012/462507/1/POLJURI_NT(2012)462507_EN.pdf)> accessed 4 April 2013.

⁶ European Commission, ‘Common European Sales Law to boost trade and expand consumer choice’ (Press Release 11 October 2011) IP/11/1175.

⁷ See Thomas Wilhelmsson, ‘The Abuse of the ‘Confident Consumer’ as a Justification for EC Consumer Law’ (2004) 27(3) *Journal of Consumer Policy* 317.

⁸ For a consideration of what influences consumers, see Rossella Incardona and Cristina Poncibò, ‘The average consumer, the unfair commercial practices directive, and the cognitive revolution’ (2007) 30(1) *Journal of Consumer Policy* 21, 21-38.

competition and choice. But at least in the immediacy, consumer confidence will be slow to react. It is only through positive experiences over time that consumers will come to know of the true benefits of the CESL.

Accordingly, this essay will assess three contestations against the CESL. Part one considers the aim in the context of problems facing Union trade generally. Part two then considers the optional nature of the CESL addressing the mechanical problems with its implementation. Finally, part three will look at the cost. Each section will be considered in terms of the potential effect on cross border sales and consumer confidence. Ultimately this essay will propose that while the CESL is a step in the right direction, it falls short of what is necessary to create a coherent Consumer Contract *acquis*, and to achieve a high standard of consumer protection across Europe.

II. A Misguided Venture

The Commission drafted the CESL in response to what they perceived to be some existing problems with the state of the Union's market principles. As it stands, only "1 in 10 Union traders export within the Union"⁹ and the Commission makes it clear that, in its opinion, contractual differences are the major obstacle facing traders contemplating cross border transactions. It claims the CESL would "improve the establishment and functioning of the internal market by facilitating the expansion of cross border trade".¹⁰ This approach over simplifies the problems that exist within the Contract law *acquis* and it becomes quickly apparent from simple reflection on the Commission's own referenced findings that diversity of contract law is one of a multiplicity of problems with regard to B2C transactions.

⁹ European Commission, 'Proposal for a Regulation of the European Parliament and of the Council on a Common European Sales Law' COM(2011) 635 final.

¹⁰ *ibid.*

The Commission has underestimated the scale of the obstacles consumers face and overreacted to their assumptions. They have underestimated in the sense that their focus on Contract law is misplaced. In practice it seems that consumers are more concerned with other issues, for example: fraud, delivery, language and the right to redress.¹¹ They have over reacted to their own claims that 44 per cent of Europeans will not buy abroad because they are “uncertain about their rights in cross border situations”.¹² They have jumped to the conclusion that “rights” are synonymous with contractual rights, but in reality this kind of aggregating “assumption leads the analysis astray”.¹³ Consumers have a high level of variance, and without a deeper understanding of what “rights” the average consumer is referring to, solutions will remain elusive.¹⁴ In fact, the Commission’s own example¹⁵ demonstrates that the level of harmonisation that already exists¹⁶ is adequate to address the

¹¹ Which, ‘Consultation Response: Response to the Green Paper from the Commission on policy options for progress towards a European Contract Law for consumers and business’ (2011) <<http://www.which.co.uk/documents/pdf/policy-options-for-progress-towards-a-european-contract-law-for-consumers-and-business-which-response-243720.pdf>> accessed 4 April 2013.

¹² European Commission Memo, ‘Attitudes towards cross-border trade and consumer protection - Analytical Report’(2011) Flash Eurobarometer 299 - The Gallop Organisation <http://ec.europa.eu/public_opinion/flash/fl_299_en.pdf> accessed 4 April 2013.

¹³ Richard Epstein, ‘Harmonization, heterogeneity and regulation: CESL, the lost opportunity for constructive harmonization’ (2013) 50 *Common Market Law Review* 207.

¹⁴ Richard Epstein, ‘Inside the Coasean Firm: Why Variations in Competence and Taste Matters’ (2012) 54(4)*Journal of Law and Economics* S41.

¹⁵ The example included a Finnish woman who did not know her legal guarantee rights when shopping in France, see the Feasibility Study <http://ec.europa.eu/justice/newsroom/news/20111011_en.htm> accessed 10 April 2013.

¹⁶ As a result of Council Directive 1999/44/EC of the European Parliament and of the Council of 25 May 1999 on certain aspects of the sale of consumer goods and associated guarantees [1999] OJ L171/12.

majority of Consumer issues.¹⁷ The real barrier for consumers is not a lack of rights in the EU market, but a severe lack of understanding and education as to what those rights are and how they apply.¹⁸ Therefore, if the CESL was to have the desired effect it would have to be accompanied by a high scale initiative to educate European consumers, otherwise the CESL may end up being overlooked and unforgivably outside the scope of consumer knowledge. The situation is thus unlikely to improve.

However, the Commission does concede that there are a multitude of problems.¹⁹ In comparison to some of these, Contract law is a “clearly identifiable obstacle to the European single market”²⁰ and can be acted on. Some obstacles, such as linguistic diversity and cultural differences are practically insurmountable. Thus, it is important not to pigeon hole the CESL as useless, improving the consistency of contract law is a good first step towards combatting many of the other issues.²¹ Seen as a first step towards further Union development, the CESL could lead to improvements in consumer confidence and cross border transactions through the union of contract law; however in its current state it is likely to be limited in its scope and any benefits to consumers may be compromised by the requirement that both parties must choose to opt into the agreement.

III. The danger of Optionality to the success of the CESL

Article 3 of the proposed regulation highlights the optional nature of the CESL. The contract mechanism will be available for parties to opt into, but only if they *agree* to do

¹⁷ Ursula Pahl, ‘The Common European Sales Law – Have the right choices been made? A Consumer Policy Perspective’ [2012] Maastricht Journal of European and Comparative Law 180.

¹⁸ European Commission, ‘Consumers’ awareness and skills worryingly low, survey finds’ (Press Release 11 April 2011) IP/11/455.

¹⁹ Directorate-General For Internal Policies (n 5).

²⁰ *ibid.*

²¹ For example, confidence in foreign markets and the problems of enforcement.

so. Thus, the CESL operates as an alternative to the contract laws of domestic states of the parties or the transaction.²² The Commission reasoned that delivering the CESL as an optional mechanism would be the “most proportionate action”.²³ It was the belief of the Commission that, being optional, the CESL would reduce the cost to traders since it would mean that those who are content trading domestically and those who wished to trade in the EU very rarely, would not face substantial legal costs in adapting their contract law processes to the new system. Upon deeper analysis, I will explain why the Commission was short sighted and the proposed ‘optional’ nature of the CESL threatens the document’s integrity as a useful contracting alternative. In turn this threatens, in particular, the advantages that the CESL will offer to the consumers through its availability and may even have a negative impact on consumers in some Member States.

The primary problem with the optional status of the CESL is that in many situations it creates a “contract of adhesion”.²⁴ That is to say, there is such a strong imbalance in contracting power between the two parties that there is an implication there has been no real bargaining over the contractual terms. This is clearly true in the case of B2C transactions, and may also be the case in Business-to-Business (B2B) transactions where there is a disproportionate difference in size between the traders. In almost all consumer transactions there is generally no opportunity for the consumer to negotiate the terms of the contract. Be it a high street or an online purchase, traders tend to offer consumers an ultimatum; either comply with their terms or take their business elsewhere. This has some incidental effects on the consumer’s relationship with the CESL and is the primary

²² Giesela Rühl, ‘Consumer Protection in Choice of Law’ (2011) 44 *Cornell International Law Journal* 569.

²³ European Commission (n 9).

²⁴ Pachl (n 17) 186-192.

cause of dissatisfaction against the proposal.²⁵ The proposal itself attempts to mitigate this issue to some degree, namely through the requirements of Article 8, which requires that the consumer must make an “explicit statement” agreeing to use the CESL.²⁶ This offers very little practical value, however it may ensure the consumer is aware the CESL applies. Nonetheless, their only two options remain to trade on these terms or to abandon the transaction. As Epstein points out, this relationship means “the likelihood of businesses offering contracts under the CESL depends on the standard of protection provided by the instrument”.²⁷ This will have a detrimental effect for consumers and may in turn decrease their confidence in the Union market if consumers fear the terms of the CESL shall be drawn against them in cross-border transactions.

In fact, in the establishment of an optional CESL lies a pertinent “conceptual dilemma”.²⁸ Since we have already acknowledged that the decision lies in the hands of the business and outside of the consumer’s control, the CESL will become either redundant or a weapon against consumers. If the optional instrument were to be established on a relatively high level of consumer protection in comparison with the national laws of the member states, it would be considered undesirable by businesses and they would not use it. But if the CESL were to be based on a level of protection considered average among the member’s states, it would be inevitably weaker than some of the domestic laws within the Union. This would lead to “a race to the bottom”.²⁹ That is to say businesses could manipulate the CESL as a tool to reduce consumer rights in countries where the national law offers a higher level of protection. This would have a detrimental effect on consumer law within the

²⁵ Susan Chae and Paul Heidhues, ‘Buyers’ Alliances for Bargaining Power’ (2004) 13(4) *Journal of Economics & Management Strategy* 731.

²⁶ European Commission (n 9).

²⁷ Epstein, ‘Harmonization, heterogeneity and regulation’ (n 13).

²⁸ Pachl (n 17) 186-192.

²⁹ *ibid.*

Member State. Such an effect may make businesses more willing to trade across borders, because they would be able to rely on the fact that the CESL would be the upper limit of consumer protection and this may improve their confidence when dealing with consumers abroad. However such a propensity is likely to be short lived, since this effect would almost certainly decrease consumer confidence in the Union market, especially in the Member States where they see their rights decreased by the advent of the CESL. Furthermore, the market becomes much less attractive if the consumer feels like the EU's regulation works in favour of the already grossly overpowered businesses.

Problematically, the optional nature of the CESL puts the mechanism out of reach of those it was designed to assist the most. It has the potential to weaken the position of consumers even further and make them reluctant to trade on the terms of the CESL. However, it may have some indirect positive effects for consumers, by empowering traders in foreign jurisdictions to undertake more cross border transactions and accordingly create a more competitive market and in theory reducing prices and increasing consumer choice. This effect would be somewhat mitigated should the cost of implementing the CESL be passed on to the consumers.

IV. The Cost of Implementation

In creating the CESL the Commission's intentions are "economic, rather than social or political".³⁰ They claim that cross border transactions are costly in legal discovery fees and adaptation,³¹ priced roughly at EUR €15,000 for entering each new jurisdiction.³² These fees are highly dissuasive for

³⁰ Simon Whittaker, 'Identifying the Legal Costs of Operation of the Common European Sales Law' (2013) 50 *Common Market Law Review* 85.

³¹ European Commission (n 9).

³² Directorate General For Internal Policies, 'Common European Sales Law: A Practical View' (2012) <[http://www.europarl.europa.eu/RegData/etudes/note/join/2012/462474/1POLJURI_NT\(2012\)462474_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/note/join/2012/462474/1POLJURI_NT(2012)462474_EN.pdf)> accessed 4 April 2013.

small traders considering entering new markets and in turn, this has a negative effect on the choice and competition available to European consumers. The CESL purports to eliminate this matter. However it seems that this issue is exaggerated and in fact to the contrary of the Commissions claims, 82 per cent of traders said that the transaction costs had no great impact upon their decision to trade across borders.³³ Thus, even if we accept the premise, it is certainly not axiomatic that the CESL will increase cross border transactions in 82 per cent of the cases.

Problematically, there should be concern that the CESL will in fact increase costs for the Union traders. The Commission mistakenly sees the CESL as the alternative regime, but realistically it is the “28th regime”.³⁴ It does not replace the contractual laws of Member States but is an optional alternative. Thus, traders have to be “well advised” on another alternative contract law system.³⁵ This could cause particular problems where relationships already exist between traders. Each party will have to make an initial investment in legal discovery costs and in the event that one party discovers they are likely to benefit from the scheme they will push its application on the trading relationship.³⁶ This process of negotiation will be costly for both sides even if enquiries have no fruition and it may be potentially destructive where there is disagreement.

Despite this, it is arguable that an enquiry into the CESL is an investment, since traders will be able to use the knowledge they gain through their one time educational enquiries to “serve a bigger market on the basis of the CESL”

³³ European Commission (n 2) p20.

³⁴ Reflection Group on the Future of the EU 2030, 'Project Europe 2030: Challenges and Opportunities' (2010) <http://www.consilium.europa.eu/uedocs/cms_data/librairie/PDF/QC3210249ENC.pdf> accessed 4 April 2013.

³⁵ Kornet (n 3) 10.

³⁶ Eric Posner, 'The Questionable Basis of the Common European Sales Law: The Role of an Optional Instrument in Jurisdictional Competition' (2013) 50 *Common Market Law Review* 261.

rather than studying other national laws.³⁷ But such uniform confidence will take time to develop, since in the beginning it will be “shrouded in uncertainty until significant jurisprudence has sprouted”.³⁸

Paradoxically, there is a conflict between the level of consumer confidence and protection the CESL offers, and its goal to increase cross border transactions.³⁹ The extent to which the proposed regulation protects consumers is broad⁴⁰ and this may have two devastating effects on the utilisation of the CESL. Firstly, maintaining a high standard of protection will discourage parties from deploying the CESL where it imposes on them higher requirements than the national laws do. Secondly, there is a danger that the CESL will decrease consumer confidence *and* dissuade cross border transactions simultaneously. The high standard of protection will mean that *if* a trader does choose to operate in a jurisdiction under the CESL, then they have to offer higher protection to consumers and “consumer protection comes at a price”.⁴¹ This cost will then be transferred onto the consumer in the form of higher sales prices.⁴² Primarily this acts like a forced warranty,⁴³ but it also means that cross border products are likely to be more expensive than those traded under domestic law. In this respect, the CESL may make cross border transactions less attractive to traders since it puts them at a competitive disadvantage. Similarly the consumer may

³⁷ Chantal Mak, 'Unweaving the CESL: Legal-economic reason and institutional imagination in European contract law' (2013) 50 *Common Market Law Review* 277, 283-5.

³⁸ Posner (n 36).

³⁹ *ibid.*

⁴⁰ Oren Bar-Gill and Omri Ben-Shahar, 'Regulatory Techniques in Consumer Protection: A Critique of European Consumer Contract Law' (2013) 50 *Common Market Law Review* 109.

⁴¹ *ibid.*

⁴² Steven Shavell, *Foundations of Economic Analysis of Law* (Harvard University Press 2004).

⁴³ Additional cost for higher protection is akin to an optional warranty; consumers usually prefer the cheaper option when all other things are similar, and that cheaper option is likely to come with an 'optional warranty' that will strengthen their consumer protection.

feel that the CESL legally forces upon them a warranty they would not have otherwise elected for.

V. Conclusions

Consumer confidence and cross border trade will increase if the consumer has stronger rights, and is aware of that fact. The CESL is a step in the right direction, but where the rights it grants are effective, the optional nature of the instrument puts it out of reach of the consumers. Perhaps more problematically, even if the CESL was available to a consumer, we know from past experience that they are unlikely to be aware of their rights. The CESL is likely to become the next generation of EU legislation unknown to consumers.

Of more concern, the CESL is of particular worth to traders where it decreases the level of consumer protection, potentially increasing trade but at a cost to the consumer. Yet, there may still be hope as the CESL offers a legitimate advantage for some businesses. It is an “indication to consumers that they are willing to grant a very high standard of protection”.⁴⁴ This is an attribute particularly appealing to traders who offer high end, less frequented and more expensive products. If the CESL were to be accompanied by a symbol, it would be a visual incentive to consumers, associated with an “excellent level of consumer protection”.⁴⁵ Through this mechanism the CESL may influence cross border trade, and in time, increase consumer confidence as domestic markets become accustomed to its presence.

Unfortunately, the CESL in its current form has too many inconsistencies and loopholes to have a positive effect on the Consumer Contract Acquis. The advantages it offers are undercut by other issues at the heart of cross-border trading and are overshadowed by its conceivable exploitation. The harmonisation of consumer law has been one of the EU’s

⁴⁴ Directorate-General For Internal Policies (n 5) 9-10.

⁴⁵ *ibid.*

greatest successes,⁴⁶ and it would be a shame to threaten this with an experimental new tool that poses significant risk and little potential advantage to consumers and traders alike.

⁴⁶ *ibid.*

Incoherence and policy in the law of causation: a cause for concern?

Alex Sprake

Abstract

This article reconsiders the law of causation in the context of criminal law. In so doing, it will explore some of the inconsistencies that can be found in the law at present, and seeks to propose an explanation as to why such inconsistencies exist. It will be seen that the apparent incoherence in the law may be explained by reference to divergent public policy influences. The approach to causation in criminal law allows judges the flexibility to decide cases on the basis of prevailing public policy; any apparent incoherence is simply the corollary of the courts' ability to decide cases on this flexible basis. Such flexibility may be seen as desirable, although it may ultimately result in legal uncertainty.

I. Introduction

In English criminal law, for an individual to be charged with a criminal offence both factual and legal causation are required: did the defendant cause the injury and are they legally the cause of that injury?⁹ Complexities arise when seeking to establish whether an individual is legally culpable for 'causing' the incident; the concept of legal causation has ironically 'caused' difficulty for the courts and academics alike. This paper will examine those difficulties, as well as the extent to which they render the current law of causation incoherent and uncertain.

II. What is legal causation?

In *Empress Car v National Rivers Authority*¹ Lord Hoffmann sought to state that causation was not a question to which a "common-sense answer" could be provided

¹ *Empress Car Co (Abertillery) Ltd v National Rivers Authority* [1999] 2 AC 22 Lord Hoffmann [2].

“without knowing the purpose and scope of the rule”.² In effect, Lord Hoffmann was highlighting the difficulties of assessing a causal influence when there is no settled and consistent doctrine of causation. Lord Hoffmann in this case indicated that there could be liability regardless of a *novus actus interveniens* (an intervening act); this effectively subverted the general understanding of the law prior to the *Empress Carcase*. Subsequently, Lord Bingham sought to restrict Lord Hoffman’s statement from having any wider application, and in so doing confined Lord Hoffmann’s comments to the facts of the *Empress Carcase*.³

The reason for Lord Bingham’s decision to limit the ambit of Lord Hoffmann’s comments is that there is a substantial difference between ensuring that companies in breach of Strict Liability environmental offences are held to account, and holding someone strictly liable for an offence which could substantially interfere with one’s liberty. Thus, by restricting Lord Hoffmann’s comments at the earliest possible point, the law avoided the rigid application of a general rule of causation, and in so doing mitigated any risk of holding a citizen strictly liable for an offence for which they are not morally to blame. Accordingly, as Lord Bingham states, “common sense answers to questions of causation will differ according to the purpose for which the question is asked”.⁴ This proposition is reflected in the work of the late Glanville Williams, who stated that “when one has settled the question of but-for causation [factual causation], the furthest test to be applied to the but-

² *Empress Car Co* (n 1) Lord Hoffmann [2].

³ *R v Kennedy* [2007] UKHL 38 [15].

⁴ *ibid.*

for cause in order to qualify it for legal recognition... atest of moral reaction".⁵

Several commentators have sought to establish what legal causation is. Hart and Honoré emphasise individual agency and human beings' ability to manipulate their surroundings; accordingly, they argue a cause is "something which interferes with or intervenes in the course of events which would normally take place".⁶ Hart and Honoré add that when this occurs the responsibility for any interferences shall be legally attributed, owing to sufficient proximity,⁷ to the individual.⁸ They build on this idea and state that an individual will only be morally culpable when the cause of the 'event' is accompanied by a range of antecedent conditions. Where an abnormal condition follows, this will become the cause in place of the original human intervention.⁹ This is problematic, because, as Norrie points out, it is hard to classify what exactly is "normal" and what exactly is "abnormal".¹⁰ Hart and Honoré concede that these conditions may mean that distinctions are "drawn in different ways in one case and the same according to the context".¹¹ This paved the way for potentially excessive judicial creativity and a lack of adherence to previous caselaw where one judge has a different interpretation or belief to another. If the law on causation partially stems from Hart's theory, this

⁵ Glanville Williams, *Textbook of Criminal Law* (2nd edn, Stevens & Sons Ltd 1983) 381.

⁶ HLA Hart and T Honoré, *Causation in the Law* (Clarendon Press: Oxford 1959) 26.

⁷ Glanville Williams, *Textbook of Criminal Law* (3rd edn, Sweet & Maxwell 2012) 195.

⁸ Hart and Honoré (n 6) 31.

⁹ Hart and Honoré (n 6) 79-80.

¹⁰ Alan Norrie 'Critique of Criminal Causation' (1991) *Modern Law Review* 690.

¹¹ HLA Hart and T Honoré (n 6) 37.

might explain its incoherence. Norrie highlights this issue: “informal ‘conditions’ are subject to... change according to their context.

It is a weak, potentially unstable foundation for legal and moral judgements”.¹² Nonetheless, it is clear that Hart and Honoré’s thesis retains much plausibility, and has arguably gained authoritative status following Lord Bingham’s recital of their work in his judgement in *Kennedy*.¹³ However, as Norrie points out, if there is uncertainty as to what will constitute an ‘abnormality’, this is perhaps not a stable basis upon which to base the UK’s law on legal causation.

As the Glanville Williams statement above suggests, legal causation is concerned with whether or not the act that caused the consequence carries with it any moral stigma in the eyes of society. Norrie adds that “individual responsibility ultimately relies upon an evaluation of what is ‘normal’ in social life”¹⁴. From this we can infer that Norrie agrees that individual fault is based on morality, though his ideas are located in the context of the socio-political implications of an individual’s actions. As such, Norrie is critical of Hart and Honoré’s theory on the basis that it emphasises too heavily the finality of human agency, and suggests that their depiction fails to take account of the wider context within which individual agency operates.

III. The coherency of the current law

The English approach to the law regarding causation has recognised some situations in which the chain of causation may be broken, as noted above. Thus in

¹² Norrie (n10) 690.

¹³ *R v Kennedy* (n 3) 14.

¹⁴ Norrie (n 10) 690.

some circumstances, a defendant whomay have committed the original act will no longer be liable, if there is a *novus actus interveniens* present. A good example is *Rv Williams*¹⁵ : here, the victim's actions constituted an intervening act which was sufficient to break the chain of causation, as the victim's actions were not foreseeable.

However, in a case similar on material facts is *Rv Roberts*¹⁶ – here, the victim's actions were held to be reasonably foreseeable in the case of sexual assault rather than merely theft (as in *Rv Williams*). This illustrates the fine dividing line between when an intervening act will and will not be deemed to break the chain of causation. But different judges may view the prospect of robbery to be as sufficient 'abnormality' in the sense envisaged by Hart and Honoré. As such the law could be said to be incoherent, unpredictable and uncertain. It is therefore difficult to ascertain when a judge will and will not find a break in the chain of causation.

Nevertheless, the law has also recognised that individuals may have physical and mental peculiarities. After the case of *Rv Blaue*,¹⁷ we have what is known as the thin skull rule: under this rule it is the "policy of the law that those who use violence on other people must take their victims as they find them".¹⁸ This allows the law to prevent defendants from escaping prosecution due to an unforeseen abnormality of the victim. It can be seen from the preceding quotation that the courts will use legal causation in order to achieve a particular policy objective. However, whether *Blaue* was a sound decision or not is up for debate. For example, if the decision in *Blaue* is scrutinised in the context of Hart and Honoré's theory, it could be argued that the attack in

¹⁵ *R v Williams & Davis* [1992] Crim LR 198.

¹⁶ *R v Roberts* [1971] EWCA Crim 4.

¹⁷ *R v Blaue* [1975] 61 Cr App Rep 271.

¹⁸ *ibid.*

Blau was an 'abnormal' cause as it was arguably not within the ordinary course of events, and the refusal to accept medical treatment due to religious beliefs is a normal cause due to the victim's deep-rooted religious beliefs. According to their theory, normal causes are not sufficient to break the chain of causation. However, in *Blau* normal causes seemingly *did* break the chain of causation. Therefore, Hart and Honoré's theory may not be as comprehensive as one may think upon first examination.

Hart and Honoré claim the victim's act did not break the chain of causation despite the refusal of medical treatment being deliberate as the victim's beliefs were involuntary. It is debateable whether the victim's decision was truly 'involuntary', especially in today's liberal society where individual rights are protected, which would logically also include a right to not believe or to abstain; though, arguably, it is the role of the judiciary to ensure protection of such rights, and this line of thought led the judges in *Blau* to hold the attacker liable for the full consequences of his action.

But questions should be raised as to how far this protection would extend. Would the courts be prepared to extend this protection to victims whomay hold deep-rooted beliefs which are based upon a 'Harry Potter'-based religion, where by a potential victim would refuse any medical treatment on the ground that they wish to be treated by magical means only? Why should this type of situation be treated differently to the situation in *Blau*?¹⁹ A possible answer to this dilemma would be that when judges make decisions, they oftenseek to make the best decision possible and for the best interests of the society they serve.

¹⁹ This example is extreme. Nonetheless, Norrie illustrates a similar argument in relation to members of Religious Sects: see Norrie (n 10) 696. Some religious sects are viewed with hostility in society. Would this societal reaction result in the courts adopting a different approach to that adopted in *Blau*?

It would seem highly unlikely that such a judge would believe that a Harry Potter religion would require the same level of protection as a Jehovah's Witness, for example. By contrast, society, or perhaps even the law, may not recognise a 'Harry Potter' religion as an actual religion. Therefore, Hart and Honoré's depiction of 'abnormality' is problematic. As noted above, the notion of 'abnormal causes' places too much emphasis on human intervention and human capacity, which accordingly fail to take account of policy objectives as well as the interests of society.

A further inconsistency which becomes evident when there is reasoning in *R v Williams* is contrasted with the reasoning in *Blaue*: why was the victim's own act in *Williams* sufficient to break the chain of causation, yet in *Blaue*, the victim's own freely informed decision was not sufficient to break the chain of causation? If the victim in *Williams* had suffered from a severe anxiety condition, would the court have been prepared to make the defendant take their victim as he found her if her reaction was partially resultant of her anxiety disorder? It is difficult to ascertain whether or not a judge would do so. It appears there are problems with the application of causation; thus, perhaps Lord Hoffman was correct to say that "one cannot give a commonsense answer to questions of causation..." as mentioned above. After all, if there is no consistent application of the rules, there cannot be any commonsense answers.

Decisions concerning medical negligence that 'cause' death also highlight largely inconsistent areas of law. In *R v Jordan*²⁰ "palpably wrong" medical treatment was held to break the chain of causation, but in *R v Smith*²¹, the judge somewhat

²⁰ *R v Jordan* [1956] 40 CR App Rep 152.

²¹ *R v Smith* [1959] 2 QB 35.

“evaded”²² the intervening act doctrine and preferred instead to adopt an ‘operating and substantial’ test for causation. Norrie also draws attention to the fact that the decision in *R v Smith* reflects policy goals, which does not support Hart and Honoré’s view, because an ‘operating and substantial’ cause makes no reference to the type of abnormality envisaged by Hart and Honoré. The particular policy goal desired here is to offer adequate protection to medical staff whom may not always have the correct facilities at their disposal. This prevents doctors from being liable because of issues relating to accessibility of medical necessities, and this also illustrates that the law will be reluctant to hold doctors liable if they lack sufficient blameworthiness.

IV. Legal causation and public policy

The contrasting approaches of Lord Hoffman in *Empress*, and Lord Bingham in *Kennedy* illustrate that in cases concerning homicide, the courts will prioritise the assessment of moral blame to ensure the culpable party is held to account,²³ whereas in pollution cases the courts prioritise the avoidance of future damage. Despite the differing approaches adopted, it is evident that both approaches demonstrate a willingness of the court to take broader socio-political considerations into account, whilst at the same time ensuring that only those who are morally blameworthy are held to be at fault for offences that can infringe upon one’s liberty and autonomy.²⁴

From this perspective, it is clear that legal causation acts as a judicial safety-

²² Norrie (n 10) 697.

²³ N Padfield, ‘Clean Water and Muddy Causation: is causation a question of law or fact, or just a way of allocating blame?’ (1995) *Criminal Law Review* 683.

²⁴ In *Kennedy* (n 3) it was held that the defendant was not the blameworthy cause, and as such had not infringed upon the victim’s autonomy.

valve', which can be used to protect society. This idea is arguably reflected by the caselaw at present, notably in the stark contrasts in approach between the medical negligence cases (particularly *R v Smith* and *R v Jordan*) and the more recent developments in strict liability offences. It seems judges will extend or limit the ambit of legal causation to protect society where society requires such an extension. This suggests that to some extent there is a coherent approach, in that judges will use legal causation to achieve a favourable outcome in the eyes of society. For example, it is clear that Lord Hoffman, in the *Empress* case, by invoking the doctrine of causation was seeking to prevent future environmental damage – an outcome undoubtedly favourable to society. By the same token, in *R v Blaue*, finding the defendant guilty despite the victim refusing treatment upheld the integrity of the victim's religious beliefs. It is arguable that deciding the case in this way was simply a reflection of prevailing public policy.

V. Conclusion

It has been shown that it is not the *outcome* of criminal causation cases that are problematic; rather, what is problematic is the *method* invoked to achieve the outcome obtained. There remains uncertainty as to the extent to which judges will feel the need to make decisions on the basis of policy. Policy considerations play an important part in determining causation; this allows judges an inherent flexibility. Such an approach is desirable, although it does mean that the current law of causation necessarily contains an element of uncertainty.

Gender, Crime and Criminology

Chris Clarke¹

Abstract

Crime is a multi-faceted problem that is driven by a wide range of means, motives and opportunities. People of all ages, education, employment, race and gender partake in criminal activities for one reason or another but no one characteristic is as prevalent in the participation of crime as that of being male. Males commit the vast majority of nearly all crimes and have, unsurprisingly, been the focus of most academic work but, in spite of this incontrovertible assertion, the reasons for this fact have been largely ignored and the study of gender and crime has been nothing more than incidental to criminological study. The purpose of this article is to examine the extent of the gender gap, the inability of pre-existing sociological theories to explain it and the new developments that may help us understand it. Thanks to insight from feminist criminologists and the response of masculinities, criminology should finally be able to break free from the dark ages of gender ignored deviance and incorporate it into the heart of criminological theory.

I. Introduction

The proposition that men commit more crimes than women is one of the few undisputed “facts” of criminology,² with men “always and everywhere more likely than women to commit criminal acts”.³ Statistically, men accounted for 76% of criminal sentences in the United Kingdom in 2011⁴ and remain a key focus of criminological debate. Unsurprisingly, the predominance of male criminality meant that most

¹ Thanks go to Professor David Gadd, Hannah Wishart, Max Melsa and Joseph Tomlinson for helpful comments on various drafts.

² Janet Lauritsen, Karen Heimer and James Lynch, 'Trends in the Gender Gap in Violent Offending: New Evidence from the National Crime Victimization Survey' [2009] 47 *Criminology* 361, 362.

³ M Gottfredson and T Hirschi, *A General Theory of Crime* (Stanford University Press 1990) 145.

⁴ Ministry of Justice, *Statistics on Women and the Criminal Justice System 2011*; A Ministry of Justice Publication under S95 of the Criminal Justice Act 1991 (Ministry of Justice 2012).

studies prior to 1980 focused on men's crime and delinquency,⁵ but the rise of feminism has encouraged criminologists to diverge from the ethnocentricity in mainstream criminology in order to formulate theories which attempt to explain why this gender gap exists.

Gender is an important consideration in the study of crime because offending is a gendered concept. Gender shapes and dictates an individual's actions by "guiding expectations and appraisals of others and the self with regard to risk-taking and criminal behaviour".⁶ A fundamental problem with the previously predominant theories of criminology was that their reliance on sex failed to consider the influence of gender on criminality. Therefore, in order to create a truly gendered understanding of crime, theory needs to illuminate how gender norms, tendencies, and differences affect individual motivations and opportunities to commit crime.

Before exploring the importance of gender and crime, it is first necessary to decouple sex from gender. Whereas the former is usually a fixed, physical characteristic,⁷ the latter is a collection of individualistic traits that dictate if an individual is masculine or feminine.⁸ Typically, sex and gender correspond, but they can differ. It is therefore important to recognise gender as a diverse range of forces, which can allow women to perform masculine-related tasks and men to perform feminine-related tasks, if we are to avoid confusing sex with gender. Criminologists must consequently

⁵ E Leonard, *Women, Crime and Society: A Critique of Criminology Theory* (Longham, 1982).

⁶ Darrel Steffensmeier, Jennifer Schawrtz and Michael Roche, 'Gender and Twenty-first-Century Corporate Crime' [2013] XX(X) *American Sociological Review* 1, 5.

⁷ Although sex is usually a biological certainty, it is not always fixed because 'one in one hundred babies are born with ambiguous genitalia' and adults can undergo sex change therapy or surgery in later life: Dana Britton, *The Gender of Crime* (Rowman& Littlefield Publishers 2011).

⁸ Suzanne Kessler and Wendy McKenna, *Gender: an Ethnomethodological Approach* (University of Chicago Press 1978).

differentiate between the two if they are to avoid creating “an empty tautology [whereby] gender collapses into sex”.⁹

II. The Gender Gap

Criminal acts are overwhelmingly carried out by men. In the UK, 90% of Class A drug offences, 88% of cases of ABH and 74% of shoplifting incidents in 2011 were committed by males.¹⁰ In general, women desist from crime more readily,¹¹ their criminal careers are shorter,¹² and they rarely participate in the highly lucrative world of organised crime.¹³ The criminal deeds perpetrated by women are less readily reliant on violence, require a higher level of provocation, and tend not to be repeated.¹⁴ Traditionally, this distinction was attributed to masculine physicality as “physical prowess and muscles are useful for committing crimes”.¹⁵ In contrast, criminologists evaluated woman’s crimes as crimes of the weak and powerless.¹⁶

Nonetheless, despite their differences, there are similarities in both gender’s patterns of offending. Most criminals, male and female, tend to be of low socioeconomic status, poorly educated, under- or unemployed, and

⁹ John Hood-Williams, ‘Gender, Masculinities and Crime: From Structures to Psyches’ [2001] 5 *Theoretical Criminology* 37, 45.

¹⁰ Ministry of Justice, *Statistics on Women and the Criminal Justice System 2011*; A Ministry of Justice Publication under S95 of the Criminal Justice Act 1991 (Ministry of Justice 2012).

¹¹ Neil Weiner, *Violent Crime, Violent Criminals* (Sage Publications 1989) 108.

¹² Britta Kyvsgaard, *The Criminal Career* (Cambridge University Press 2003) 135.

¹³ Commonwealth of Pennsylvania, *Organised Crime in Pennsylvania: The 1990* (Pennsylvania Crime Commission 1991).

¹⁴ Weiner N, *Violent Crime, Violent Criminals* (Sage Publications 1989) 35.

¹⁵ Marisa Silvestri and Chris Crowther-Dowey, *Gender & Crime* (Sage Publications 2008) 52.

¹⁶ See Darrell Steffensmeier, *The Fence in the Shadow of Two Worlds* (Rowman and Littlefield 1986) and Anne Worrall, *Offending Women: Female Lawbreakers and the Criminal Justice System* (Routledge 1990).

disproportionately from minority groups.¹⁷ Additionally, both males and females tend to commit crime at a comparable rate after their first offence¹⁸ and tend to react equally to social pressures and legal forces.¹⁹ While these similarities are important to recognise, they cannot compensate for the aforementioned differences, which show that gender is a crucial aspect of understanding criminal behaviour.

The gender gap is a universal characteristic of criminal statistics, persisting around the world regardless of how social and economic conditions differ,²⁰ and it is therefore essential that criminologists “recognise gender as a social construct and not simply as a statistical variable”.²¹

However, these statistics must be treated with caution as the world reacts differently to each gender’s offending, taking into account the characteristics of both crime and punishment. For example, women shoplifters are sent “back into the private sphere of the home, while more professional male thieves [are turned] into the hands of the public police and courts”²² and because of this chivalrous bias, women are systemically under-accounted for in crime statistics. Instead, they are left to the informal control of the family whilst men feel the full force of the law.

¹⁷ Darrell Steffensmeier and Emilie Allan, ‘Gender and Crime: Toward a Gendered Theory of Female Offending’ [1996] 22 *Annual Review of Sociology* 459, 465.

¹⁸ Deborah Denno, ‘Gender, Crime, and the Criminal Law Defenses’ [1994] 85 *The Journal of Criminal Law and Criminology* 80, 94.

¹⁹ H Boritch and J Hagan, ‘A century of crime in Toronto: gender, class and patterns of social control 1859-1955’ [1990] 28 *Criminology* 567, 601.

²⁰ Darrell Steffensmeier and Emilie Allan, ‘Gender and Crime: Toward a Gendered Theory of Female Offending’ [1996] 22 *Annual Review of Sociology* 459, 465.

²¹ Frances Heidensohn and Marisa Silvestri, ‘Gender and Crime’ in Mike Maguire, Rod Morgan and Robert Reiner, *The Oxford Handbook of Criminology* (5th edn, Oxford University Press 2012) 336, 338.

²² Hagan J, Simpson J and Gillis A, ‘The Sexual Stratification of Social Control: A Gender-Based Perspective on Crime and Delinquency’ [1979] 30 *The British Journal of Sociology* 25, 26.

III. Gender Blind Criminology

Although criminologists have tried admirably to provide general explanations of crime and deviance, their theories were designed to explain male, rather than female, crime.²³ None of the most widely accepted criminological theories adequately account for the influence of gender, and in doing so have “resisted this obvious insight with energy comparable to that of medieval churchmen denying Galileo”.²⁴

Merton’s theory of strain and anomie²⁵ suggests that unequal social structures and opportunities to attain success lead to crime. As such, the marginalisation and victimisation of the female population should make them more criminally active than their male counterparts when this is clearly not the case. Durkheim,²⁶ along with Broidy&Agnew,²⁷ explains this apparent discrepancy as a product of women’s distinct emotional responses to pressure and strain, but this alteration to Merton’s theory is unsatisfactorily ad-hoc and reactionary.

Similarly, Hirschi’s control theory²⁸ suggests that inhibiting each individual’s desire to deviate can reduce crime. However, by focusing upon male criminality, rather than female conformity, it fails to explain why females more readily follow expected patterns of attitude and behavior, ignoring the importance of gender expectations in criminal careers.

²³ See Dana Britton, *The Gender of Crime* (Rowman& Littlefield Publishers, Plymouth 2011); E Leonard, *Women, Crime and Society: A Critique of Criminology Theory* (Longham, 1982); Frances Heidensohn, ‘The Deviance of Women: A Critique and an Enquiry’ [1968] 19 British Journal of Sociology 160.

²⁴ Frances Heidensohn, *Women and Crime* (New York University Press 1987) 22.

²⁵ Robert Merton, ‘Social structure and anomie’ [1938] American Sociological Review 672.

²⁶ Emile Durkheim, *Suicide: A Study in Sociology* (Simon and Schuster 1951).

²⁷ Lisa Broidy and Robert Agnew, ‘Gender and Crime: A General Strain Theory Perspective’ [1997] Journal of Research in Crime and Delinquency 275.

²⁸ Travis Hirschi, *Causes of Delinquency* (University of California Press 1969).

Additionally, labelling theory suggests that “social groups create deviance by making rules whose infraction creates deviance”²⁹ but it is systemically incapable of explaining the gender gap. Generally, men find themselves in more powerful positions than women and as such, should be able to shape the social rules that define deviance. Therefore, women should be labelled more criminal than men because of their inability to influence and thereby abide by the *rules*. The fact that women are not more but less likely to be labelled deviant than men illustrates the futility of the theory when it comes to explaining the gender gap.

Finally, the gender equality hypothesis was formulated, which predicted that increased gender equality would lead to an increase in female crime.³⁰ The emergence of the “mean girl”³¹ and “girl gang”³² was attributed to women’s liberation and the consequences of being “freed from the constraints of their gender”.³³ However, in the United States of America, between 1965 and 2004 the female share of arrests only rose from 11% to 23% in 2004,³⁴ despite the fact “women have substantially gained on men throughout the wage distribution over the last four decades”.³⁵ Additionally, violent and serious property crime remained an overwhelming male activity and

²⁹ Howard Becker *Outsiders: Studies in the Sociology of Deviance* (The Free Press 1963) 9.

³⁰ Freda Adler, *Sisters in Crime* (McGraw-Hill 1975); Rita Simon, *Women and Crime* (Lexington Books 1975).

³¹ Freda Alder and Anne Worrall, *Girl's Violence: Myths and Realities* (SUNY Press 2004); Jessica Ringrose, ‘A new Universal Mean Girl: Examining the Discursive Construction and Social Regulation of a New Feminine Pathology’ [2006] 16 *Feminism and Psychology* 405.

³² Anne Campbell, *The Girls in the Gang* (2nd edn, Blackwell 1991).

³³ Meda Chesney-Lind, ‘Patriarchy, Crime, and Justice: Feminist Criminology in an Era of Backlash’ [2006] *Feminist Criminology* 6, 11.

³⁴ Jennifer Schwartz and Darrell Steffensmeier, ‘The Nature of Female Offending: Patterns and Explanation’ in Zaplin R, *Female Offenders: Critical Perspective and Effective Interventions* (Jones & Bartlett 2007) 43, 48.

³⁵ This is based on weekly wages for men and women recorded between 1963 and 2005- David Autor, Lawrence Katz and Melissa Kearney, ‘Trends in US Wage Inequality: Revising the Revisionists’ [2005] 90 *The Review of Economics and Statistics* 302.

increased equality decreased the number of women being pushed into offending by oppressive patriarchy.³⁶

Ultimately, although these socially grounded theories of crime represent an improvement on Lombroso's crude biologism,³⁷ they all represent an unnecessary detour that stalled efforts to develop a gendered theory of crime by focusing on macro issues that failed to detail micro matters at the heart of the debate.

IV. Insights from Feminist Criminology

In an attempt to combat the inability of traditional theories to explain the lack of female crime, feminists have argued against the theories of crime that attribute women's lower crime rates to "their domesticity, maternal instinct, and passivity"³⁸ or to their "sexual coldness, weakness, and undeveloped intelligence".³⁹ By challenging biological and sociological stereotypes, feminists questioned the overly simplistic assumptions of primarily male-asserted theories in an attempt to cultivate a more gender aware approach to criminology that reflects the true influence of masculinity and femininity.⁴⁰

In general, society has different standards and expectations for girls and boys, men and women. Whereas boys are often seen as active and adventurous exhibitionists, girls are seen as mature and cunning bourgeois. Whereas the difference between what is boisterous and what is criminal is blurred, the difference between what is feminine and what is

³⁶ Darrell Steffensmeier & Jennifer Schwartz, 'Trends in Female Criminality: Is Crime Still a Man's World?' in Price and Sokoloff, *The Criminal Justice System and Women: Offenders, Prisoners, Victims, and Workers* (3rd edn, McGraw Hill 2004) 95-111.

³⁷ Cesare Lombroso and William Ferrero, *The Female Offender* (Hein 1895).

³⁸ Carol Smart, *Women Crime and Criminology* (Routledge 1976) xiv.

³⁹ Cesare Lombroso & Guglielmo Ferrero, *The Female Offender* (Hein 1895) 151.

⁴⁰ Joanne Belknap, *The invisible woman: gender, crime, and justice* (2nd edn, Michigan University Press 2001); Maureen Cain, 'Towards Transgression: New Directions in Feminist Criminology' [1990] 18 *International Journal of Sociology of Laws* 1-18.

criminal is clear. As a result, boys are policed more for legitimate criminal activities and girls are policed more for gender-inappropriate deviance- reflecting broad “social and political concerns to “police” girls in social life and to reinforce gender stereotypes”.⁴¹

Rather than controlling their actions by criminalisation, society controls female actions through suppression, as an extension of male patriarchy and female subordination. Radical feminists see “patriarchy as [an] unremittingly oppressive tyranny driven by the ‘masculine traits’ of aggression, intimidation and physical violence”,⁴² and believe men ultimately maintain their power through the fear of rape.⁴³ Nowadays, few hold such a radical view, but the feminist position still maintains that society ensures female subordination in more subtle ways. For example, as the moral panic of the female offender has emerged, female imprisonment has grown and the “courts are [now] imposing more severe sentences on women for less serious offences”.⁴⁴

Whereas traditional paternalist tendencies have perpetuated the myth that women are spared the cruelty of prison, statistics show that the “number of women sentenced to immediate imprisonment in England and Wales grew faster than comparable figures for males for much of the 1990s and into the twenty first century”.⁴⁵ However, whilst feminists have worked toward the elimination of gender inequality, they are yet to offer a comprehensive alternative

⁴¹ Loraine Gelsthorpe and Anne Worrall, ‘Looking for Trouble: A Recent History of Girls, Young Women and Youth Justice’ [2009] 9 *Youth Justice* 209, 210.

⁴² Steve Hall, ‘Daubing the drudges of fury’ [2002] 6 *Theoretical Criminology* 35, 47.

⁴³ Susan Brownmiller, *Against Our Will: Men, Women and Rape* (Random House Publishing Group 1975); Susan Griffin, ‘Rape: The All American Crime’ in Cox, *Female Psychology: The Emerging Self* (Palgrave Macmillan 1971) 291-315.

⁴⁴ Home Office, *Statistics on Women and the Criminal Justice System: Section 95 Report* (Home Office 2004) 3.

⁴⁵ Frances Heidensohn and Marisa Silvestri, ‘Gender and Crime’ in Maguire, Morgan & Reiner (eds), *The Oxford Handbook of Criminology* (5th edn, Oxford University Press 2012) 336, 352.

and have actually created a bifurcated criminology,⁴⁶ whereby one side studied traditional lines of enquiry and the other side approached the field through a feminist lens—representing an another unfortunate discourse as opposing parties went about trying to replace one partisan view with another.

V. A Response to Feminism through Masculinities

As is now recognised universally, “nothing is as frequently associated with criminal behaviour as being a male”.⁴⁷ Criminology’s contemporary turn to masculinities is therefore a long overdue attempt to address the “apparently obvious, and yet seemingly so difficult to explain, maleness of crime”.⁴⁸ Viewing masculinity as a socially shaped concept, Miedzian⁴⁹ believes that the influence of the media and the military have led people to believe that violence is an acceptable form of masculine behaviour and that the link between masculinity and crime is now widely established.⁵⁰ *Thelma and Louise*, the patriarchal backlash blockbuster, is an unexpected yet educative example of a number of these concepts, illustrating: how girls can “do” masculinity; how informal social influences dictate acceptable forms of gender behaviour; and how the “miscellany of masculinity”⁵¹ structures gender relations.

⁴⁶ Laura Brown, ‘Gender-role analysis: A neglected component of psychological assessment’ [1986] 23 *Psychotherapy: Theory, Research, Practice, Training* 243.

⁴⁷ Edwin Sutherland, ‘Development of the Theory’ in Schuessler, *Edwin H Sutherland on Analyzing Crime* (University of Chicago Press 1942) 13, 19.

⁴⁸ Richard Collier, *Masculinities, Crime and Criminology* (Sage Publications 1998) 177.

⁴⁹ Myriam Miedzian, *Boys will be Boys* (Lantern Books 1991).

⁵⁰ Lee Bowker, *Masculinities and Violence* (Sage Publications 1998); James Messerschmidt, *Masculinities and Crime* (Rowman & Littlefield Publishers 1993).

⁵¹ Manohla Dargis, ‘Roads to Freedom’ [1991] 1 *Sight and Sound* 15.

West and Zimmerman⁵² describe gender as an everyday, routine accomplishment that can be achieved in both an orthodox and unorthodox fashion. *Doing gender* involves earning continuous societal recognition, and it is usually achieved through socially accepted accomplishments such as having a family, a job, and a car. However, males who are unable to express their masculinity legitimately are more likely to resort to violence or criminal activities.⁵³ Therefore, these men of typically poor backgrounds resort to crime in order to assert their toughness through actions that are “traditionally symbolic of untrammelled masculinity”.⁵⁴

In order to rationalise the concept of masculinity, Connell⁵⁵ contextualised the concepts of patriarchy, domination, and exploitation through which men are deemed powerful. In doing so, he “drew together feminist perspectives on the social construction of gender”,⁵⁶ whilst remaining “conversant with the range of troubling dilemmas and contradictions many men experience in their everyday attempts to negotiate interpersonal relationships”.⁵⁷ By adopting such an approach, Connell clarified the flexibility and plurality of masculinities and their hierarchical structures, laying the foundations for insight into crime and gender through hegemonic and subordinate masculinities.

Messerschmidt⁵⁸ built on West & Zimmerman’s and Connell’s foundations by exploring how gender is accomplished in society and why criminal behaviour is a

⁵² Candace West and Don Zimmerman, ‘Doing Gender’ [1987] *Gender & Society* 125.

⁵³ Jessie Krienert, ‘Masculinity and Crime: A Quantitative Exploration of Messerschmidt’s Hypothesis’ [2003] *Electronic Journal of Sociology* 1-25

⁵⁴ Albert Cohen, *Delinquent Boys* (Free Press 1955) 140.

⁵⁵ Raewyn Connell, *Gender and Power: Society, the Person and Sexual Politics* (Allen & Unwin 1987); *Masculinities* (Allen & Unwin 1995); *Gender* (Polity Press 2002).

⁵⁶ Frances Heidensohn and Marisa Silvestri, ‘Gender and Crime’ in Maguire, Morgan & Reiner (eds), *The Oxford Handbook of Criminology* (5th edn, Oxford University Press 2012) 336, 348.

⁵⁷ David Gadd and Stephen Farrall, ‘Criminal Careers, Desistance and Subjectivity’ [2004] 8 *Theoretical Criminology* 123, 128.

⁵⁸ James Messerschmidt, *Masculinities and Crime* (Rowman & Littlefield Publishers 1993).

resource for men to do masculinity when their economic and social resources are inadequate. Traditionally, men are expected to provide for their families but if they are unable to do this, they will seek out other “masculine validating resources”,⁵⁹ which, for a group of adolescent males, can mean having the “balls to go break into a house or steal a car [or] screwing the most girls..., holding the most beer, smoking the most weed”.⁶⁰ In keeping with the desire of all humans to fill their socially mandated role, crime is a result of the male drive to ruthlessly accomplish masculinity by any means necessary, as a result of excessive individualistic gender tension.⁶¹

Whereas traditional theories are unable to explain why “boys disproportionately participate in computer crime compared with girls”,⁶² the concept of masculinities is able to offer an explanation for the prevalence of male criminality in cybercrime. According to this theory, even in an anonymous and force-free environment of the cyber world, the gender gap remains because men are inherently more likely to offend in order to facilitate profit-making and assert their dominance over others.

However, Jefferson⁶³ saw this version of masculinities as over socialised, and, in contrast, attributed greater importance to how men interpret and react to their own psychic considerations and life history. Whereas Messerschmidt’s theory only predicted which social groups, rather than which individuals, were likely to resort to crime, he thought a psychoanalytic analysis was more accurate because it can “encompass both the messy reality of actually

⁵⁹ James Messerschmidt, *Masculinities and Crime* (Rowman & Littlefield Publishers 1993) 83.

⁶⁰ *ibid* 70.

⁶¹ Steve Hall and Simon Winlow, ‘Culture, Gender and Violence: a key problem’ [2003] 53 *Criminal Justice Matters* 14.

⁶² Byongook Moon and others, ‘Gender, General Theory of Crime and Computer Crime’ [2012] XX(X) *International Journal of Offender Therapy and Comparative Criminology* 3, 14.

⁶³ Tony Jefferson, ‘Subordinating Hegemonic Masculinity’ [2002] 6 *Theoretical Criminology* 63.

existing gender relations, the diversity of actual men and women's relationships to discourses of masculinity and femininity, and the underlying psychological processes".⁶⁴ Thus, by elevating the importance of psychoanalytic theory and incorporating it into masculinities, criminology has been able to better accommodate each individual's "unique (sexual) genesis"⁶⁵ and understand the vulnerabilities felt by even the most vicious criminals.

Nonetheless, critics still question how accurately gender can explain crime through masculinities given the following: the pervasiveness of male criminality suggests that masculinity is deeply distressed;⁶⁶ the notion that crime is a masculine trait is a tautology;⁶⁷ and the concept of hegemonic masculinity is too broad of a sociological concept.⁶⁸ Indeed, in practice, masculinity remains in its infancy and needs to be developed to encompass a more explicit reference to the psyche and clearer isolation of gender itself. However, in theory, it offers the most feasible route towards understanding the relationship between gender and crime.

Traditionally, masculinity has been identified as a propensity to fight, be tough, and dominate,⁶⁹ and invariably some men are always "reaffirming collective male violence, misogyny, and homophobia [and] are, in a sense, familiar folk devils".⁷⁰ In reality however, not all men are aggressive and the concept of masculinities now provides both macro and micro explanations as to what it is "not as working-class, not as migrants, not as underprivileged individuals, but as

⁶⁴ Tony Jefferson, 'Subordinating Hegemonic Masculinity' [2002] 6 *Theoretical Criminology* 63, 81.

⁶⁵ *ibid* 76.

⁶⁶ Anne Campbell *Men, women and aggression* (Harpercollins 1993).

⁶⁷ M Silvestri & C Crowther-Dowey, *Gender & Crime* (Sage Publications 2008) 64.

⁶⁸ John Hood-Williams, 'Gender, Masculinities and Crime: From Structures to Psyches' [2001] 5 *Theoretical Criminology* 37.

⁶⁹ Joachim Kersten, 'Crime and Masculinities in Australia, Germany and Japan' [1993] 8 *International Sociology* 461.

⁷⁰ Richard Collier, *Masculinities, Crime and Criminology* (Sage Publications 1998) 175.

men, that induces them to commit crime”.⁷¹ By exploring the sociological and psychoanalytical effects of gender, masculinity helps deepen our understanding of male behaviour and therefore crime.

VI. Conclusion

Although it is true that gender clearly helps explain crime, it is not the only socio-economic factor that influences it. When a poorly educated, black teenage boy steals an expensive smartphone from a middle class, female University student, there is a lot more to consider than simply gender. However, considering the predominance of male crime, it is important not to under-estimate the influence of gender on criminological study. Regardless of historical tendencies to overlook its importance as a matter of social coincidence, and contemporary attempts to shoehorn it into pre-existing sociological theories, it is vital to treat gender as “fundamental, not incidental”⁷² to the study of crime.

As such, developments in the study of masculinities have been a welcome addition, helping to remove the barriers created by gender specific theories and encouraging academics to ask why masculinity causes crime rather than why men commit crime. However, in order to further understand and explain how gender causes crime, it is important that criminologists adopt a more nuanced approach towards the concept of masculinity if they are to fully appreciate the gendered nature of deviance and allow criminology to finally face its own sex question.

⁷¹ Elizabeth Grosz, ‘Feminist Theory and the Challenge to Knowledge’ [1987] 10 *Women’s Studies International Forum* 6.

⁷² Dana Britton, *The Gender of Crime* (Rowman & Littlefield Publishers 2011) 141.

A Comparative Analysis of the RNR based Sex Offender Treatment Programmes

JessicaCapello Mathews

Abstract

The effectiveness of sex offender treatment programs has provoked widespread debate among academics; whilst many praise their ability to reduce re-offending and promote desirable intrinsic outcomes, others condemn their alleged potential for indifferent or even negative results. This article seeks to determine whether current theories, namely the risk, need and responsivity (RNR) theory, can consistently provide positive results in comparison with cognitive-behavioural therapy (CBT). The effectiveness of the theory of RNR in conjunction with alternative theories will be examined, before acknowledging the difficulties that may arise when attempting to conclusively assess this issue. After reviewing the evidence, it becomes apparent that RNR based CBT are the most effective treatments available, although one must acknowledge that this is currently the most widely used method and therefore there is significantly more data available to compare. At present RNR based CBT can effectively reduce re-offending and enable intrinsic benefits for a variety of sex offenders.

When examining the topic of sex offender treatment programs in general one must be aware that opinions, from the theory through to the effectiveness, lack a decided amount of consensus.¹ However at present, the theory that is most effective in providing intrinsic benefits to sex offenders, whilst at the same time providing reduced rates of re-offending, is conceptualised around the collective ideas of Andrews, Bonta and Hodge.² Their principles of risk, need and responsivity (RNR) encompass a number of strategies designed to target and correct individual traits of offenders,

¹ Dennis Howitt, *An Introduction to Forensic and Criminal Psychology* (Harlow: Pearson 2011).

² (1990) cited in DA Andrews, James Bonta & J Stephen Wormith, The Risk-Need-Responsivity (RNR) model: Does adding the Good Lives Model contribute to effective crime prevention? (2011) 38 *Crim Justice Behav* 735, 751.

which encourage that particular individual to offend.³ Reviewing the historical context, the first sex offender treatment programs were predominantly behavioural in their nature and founded on the theory of Pavlov's Classical Conditional Model, which used negative association in order to try and curb undesirable behaviours.⁴ By the late 1900s the focus had shifted to cognitive processes, which put greater emphasis on identifying and promoting change in the dysfunctional thought processes and emotions of offenders.⁵ A combination of these two processes, named cognitive-behavioural treatments (CBT), was formulated and interlinked with RNR principles to provide treatments that can be extremely effective in correcting both the undesired physical aspects of offenders, with treatments such as hormone control, and the psychological aspects such as difficulties with social skills.⁶ Although other theoretical frameworks do exist, the most positive results have derived from a combination of RNR and CBT programs.⁷ It must be noted, that some academics have gone as far not only to question the effectiveness of these sex offender treatment programs, but to argue that among certain sectors of sex offenders, sex offender treatment programs can actually be detrimental and cause degeneration in behaviour.⁸ Though these claims are largely unfounded, to ensure maximum effectiveness care must be taken when implementing treatment programs. To yield the most beneficial and effective results, sex offender treatment programs must distinguish their treatments based on differing types of sex offenders, for example rapists and pedophiles are far less likely to prosper under generic rather than individually

³ Howitt (n 1).

⁴ Aviva Moster, Dorota W Wnuk & Elizabeth L Jeglic, Cognitive Behavioural Therapy Interventions with Sex Offenders (2008) 14 *J Correct Healthcare* 109, 110.

⁵ Abel et al, 1985; Marshall, 1971 cited in Moster (n 4).

⁶ Moster (n 4).

⁷ Howitt (n 1).

⁸ Lawrence F Jones, Iatrogenic Interventions with Personality Disordered Offenders (2007) 13 *Psychol Crime Law* 69, 70.

tailored treatment programs.⁹ Finally, it is also worth questioning the definition of ‘effective’ when assessing the effectiveness of sex offender treatment programs. Should effectiveness be limited purely to whether or not any re-offending occurs, or should other benefits of treatment, such as regain of self-worth or increased motivation to change in the individual,¹⁰ be reason enough for continuation of these treatments.

Many strongly believe that the theory of RNR “should be primary consideration in the design and implementation of programs for sexual offenders”,¹¹ however other theories such as the Good Lives Model (GLM) have emerged offering alternative methods, goals and desired outcomes. Though influence may be drawn from these alternate treatments, RNR has been consistently shown to be the most effective theory basis.¹² RNR is based on the theory that treatment should focus on first assessing offenders in terms of risk and decide how intensive treatment should be in relation to the level of risk.¹³ Static risk factors, such as previous criminal convictions, cannot be changed and are good indicators of long-term risks, dynamic stable risk factors are those which can be changed with correct treatment, such as deviant sexual interests, and dynamic acute risk factors are associated with offending in the short-term, such as substance abuse.¹⁴ The ‘need’ factor consists of then highlighting the problematic characteristics of the offender and categorising them into

⁹ Marianne Barrett, Robin J Wilson & Carmen Long, Measuring Motivation to Change in Sexual Offenders From Institutional Intake to Community Treatment (2003) 15 *Sex Abuse* 269, 280.

¹⁰ Dennis MDoren & Pamela M Yates, Effectiveness of sex offender treatment for psychopathic sexual offenders (2008) 52 *Int J Offender Ther Comp Criminol* 234, 236.

¹¹ R Karl Hanson, Guy Bourgon, Leslie Helmus & Shannon Hodgson, The Principles of Effective Correctional Treatment Also Apply To Sexual Offenders.

A Meta-Analysis (2009) 36 *Crim Justice Behav* 865, 866.

¹² Howitt (n 1).

¹³ Howitt (n 1).

¹⁴ Don Grubin, Sexual offending and the treatment of sex offenders (2007) 6 *Psychiatry* 439, 440.

‘criminogenic’ and ‘non-criminogenic’ factors. Criminogenic factors are those that increase the likelihood of re-offending (such as impulsiveness), whereas non-criminogenic factors are those that do not necessarily affect offending, such as low self-esteem.¹⁵ Identifying the need is fundamental to improving effectiveness of treatment programs, as it immediately tailors treatments to the offender’s specific needs. Finally the ‘responsivity’ factor consists of calculating which methods of treatment are likely to be most effective for the individual offender.¹⁶ Combining assessment of these factors can ensure that all aspects contributing to criminal behaviour are targeted, this can significantly aid creating effective sex offender treatment programs with the goal of reaching the primary, but not exclusive, goal of crime prevention and desistance from offending.¹⁷

GLM focuses on a more positive rehabilitation model and suggests, that by improving self-worth and encouraging good relationships, sex offenders are less likely to re-offend.¹⁸ It is widely acknowledged that “GLM views RNR as incomplete with its eye fixated on the end point, the elimination of criminal behaviour”,¹⁹ rather than the personal fulfilment of offenders. Sex offenders respond more positively to treatment if they are treated firmly but with respect by their therapists,²⁰ and this is a core value of GLM. Despite the claims that RNR does not include aspects such as “therapeutic alliance”,²¹ the RNR model extensively encourages training of probation officers to build natural and positive relationships with the offenders in their care,²² in

¹⁵ D A Andrews & James Bonta, ‘A commentary on Ward and Stewart’s model of human needs’ (2003) 9 *Psychol Crime Law* 215, 217.

¹⁶ Howitt (n 1).

¹⁷ Andrews (n 2).

¹⁸ Tony Ward & Claire Stewart, (2003). *Criminogenic Needs and Human Needs: A Theoretical Model*, (2003) 9 *Psychol Crime Law* 125, 126.

¹⁹ Andrews (n 2) 740.

²⁰ Drapeau (2005) cited in Howitt (n 1).

²¹ Andrews (n 2).

²² Andrews (n 2).

order to promote successful outcomes after treatment. This illustrates that misconceptions of RNR practices are largely incorrect and when considering theories most suited for application to sex offender treatment programs, the RNR model should be paramount.

Critics of RNR would also argue that it dedicates little time and attention to clinical issues such as motivation of sex offenders.²³ The belief that RNR does not concentrate sufficiently on these issues, is something that is largely a consequence of agencies incorrectly applying RNR principles or not adequately incorporating clinical aspects, rather than the issues not being included.²⁴ Whilst GLM may offer assistance in terms of aiding in the treatment of motivational issues, there is a danger of focusing primarily on creating a fulfilling life, in that crime prevention becomes a secondary goal.²⁵ The aim of treatments that follow the RNR principles is to combine protecting the community (by dedicating focus to criminogenic needs) and nurturing the needs of the offender.²⁶ While RNR principles provide an active model for desistance of offending, and following these principles can result in strong positive crime prevention effects;²⁷ the RNR theory also includes at the very least a basic understanding and incorporation of GLM principles on the majority of topics. Furthermore RNR's effectiveness is illustrated in practice; on the one hand RNR treatments lead to considerable reductions in re-offending whereas other methods, such as GLM 'yield minimal results'.²⁸ The RNR model encourages not only support of the psychological and emotional needs of sex offenders, but also stresses the importance of crime prevention. By combining these factors RNR has yielded more positive results than any other

²³ Andrews (n 2).

²⁴ Andrews (n 2).

²⁵ Andrews (n 2).

²⁶ Moster (n 9).

²⁷ Bonta et al, (2006) cited in Karen D'Silva, Conor Duggan&Lucy McCarthy, Does Treatment Really Make Psychopaths Worse? A Review of the Evidence (2004) 18 *JournPersDisord* 163.

²⁸ Andrews (n 2) 736.

model²⁹ demonstrating its superiority on providing effective sex offender treatment programs.

In assessing the effectiveness of sex offender treatment programmes in general, the difficulties that arise in terms of the validity of results, must be considered. There is at least consensus among most that many past treatments have not been particularly successful in reducing re-offending,³⁰ the real debate is in reviewing more current programs. Despite positive results, poor strategies or bias in investigations can invalidate findings that at first appear to reduce re-offending significantly. There are many conflicting theories as to the success of sex offender treatment programs in practice, due to the fact that when evaluating successes or failures most academics have differing views of which studies are adequate or the 'best' to use.³¹ Kenworthy et al³² included data collected from studies that used self-reports in terms of change in psychological characteristics, however Rice and Harris believed that their results were not valid.³³ It is understandable, though not necessarily correct, that Rice and Harris concluded that sex offender treatment programs show no overall treatment affect, because in their opinion, the studies were simply not to satisfactory standard.³⁴ Although RNR treatment programmes do demonstrate positive results, one must at least be aware that difficulties with the validity of results do exist.

Difficulty also arises when examining uncontrollable factors and the effects they may have, which is illustrated in

²⁹ Andrews (n 2) 736.

³⁰ R Karl Hanson, Arthur Gordon, Andrew J R Harris, Janice K Marques, William Murphy, Vernon L Quinsey, & Michael C Seto, 'First report of the Collaborative Outcome Data Project on the effectiveness of psychological treatment of sex offenders' (2002) 14 Sex Abuse 169, 170.

³¹ Anthony R Beech & Leigh Harkins, 'Examining the Impact of Mixing Child Molesters and Rapists in Group-Based Cognitive-Behavioural Treatment for Sexual Offenders' (2008) 52 Int J Offender Ther Comp. Criminol 37.

³² (2004) cited in Beech (n 31).

³³ Marnie E Rice, & Grant T Harris, 'The size and sign of treatment effects in sex offender therapy' (2003) 989 Ann N Y AcadSci 428, 438.

³⁴ Rice (n 33).

Schmucker and Lösel's results.³⁵ When initially interpreting and evaluating the results of their study, both surgical and chemical castration appeared to yield the most significant decrease in re-offending for sex offenders in general. On further analysis of other potential influencing factors attributing to the positive results of surgical castration, it became evident that offender characteristics may have had a large affect. It is not unreasonable to assume that those willing to undertake intrusive, non-reversible surgery may be the most highly motivated to reform from the outset,³⁶ and that this motivation at least partially contributed to the low re-offending statistics. Similarly chemical castration, which is a reversible procedure, originally showed re-offending levels significantly diminish. However, after the discontinuation of treatment, hormone levels within the patient returned to normal and further studies have demonstrated that consequently re-offending also increases.³⁷ It becomes clear that although sex offender treatment programmes can be shown to be effective, there are difficulties such as the influence of external factors, which cannot be ignored.

This is not to say that sex offender treatment programs are entirely ineffective at decreasing the likelihood of re-offending. On the contrary, in terms of significantly diminishing the level of re-offending, programs that follow CBT techniques have appeared consistently effective in the majority of studies.³⁸ Results of RNR based CBT programs show re-offending rates after a 5-year follow up are between 5-10% lower in treated sex offenders, than untreated sex offenders.³⁹ These techniques are more effective in practice and "superior to purely behavioural treatments",⁴⁰ due to a number of reasons. They are typically highly structured and

³⁵ Martin Schmucker & Friedrich Lösel, 'Does sexual offender treatment work? A systematic review of outcome evaluations' (2008) 20 *Psicothema* 10, 16.

³⁶ Schmucker (n 35).

³⁷ Berlin & Meinecke (1981) cited in Schmucker (n 35).

³⁸ Howitt (n 1).

³⁹ Duwe & Goldman, (2009) cited in Howitt (n 1).

⁴⁰ Hanson (n 30).

intensive (sometimes involving in excess of 300 hours over 28 weeks),⁴¹ and include specific cognitive treatments such as anger control and interpersonal problem solving, which is known to aid positive outcomes.⁴² An example of a treatment program that incorporates CBT is the Sex Offender Treatment Program (SOTP), which was set up in 1991.⁴³ It is divided into four main sections; core treatment, extended treatment, booster treatment and thinking skills.⁴⁴ While extended treatments focus more on the behavioural aspects of offending, the core treatments are the most extensive and focus around cognitive factors, for example making the offender aware of what dynamic factors such as mood can increase their level of risk on a day-to-day basis.⁴⁵ This again relates back to RNR principles, illustrating the effectiveness of identifying and treating risk factors detrimental to positive results.

In a study of SOTP, Beech et al found that after psychological testing nearly all the dynamic risk factors examined had improved, further demonstrating the effectiveness of CBT in practice.⁴⁶ Additionally, SOTP and CBT promote techniques such as group therapy;⁴⁷ in taking part offenders are exposed to less extreme distortions, which can directly reduce their own distortions.⁴⁸ It is at first difficult to assess if any treatments actually cause a significant decrease in re-offending among sex offenders, however programs that incorporate CBT that follow RNR principles

⁴¹ Polaschek et al, (2004) cited in Howitt (n 1).

⁴² Mark W Lipsey, Nana A Landenberger & Sandra J Wilson, 'Effects of Cognitive Behavioural Programs for Criminal Offenders' (2007) 6 Campbell Systematic Reviews 1, 5.

⁴³ Howitt (n 1).

⁴⁴ Anthony Beech, Richard C Beckett & Dawn Fisher, '*Step 3: An evaluation of the Prison Sex Offender Treatment Programme* (London: Stationary Office 1998).

⁴⁵ Beech (n 44).

⁴⁶ Beech (n 44).

⁴⁷ Howitt (n 1).

⁴⁸ Brian P Marx, Robert Miranda Jr & Lori A Meyerson, 'Cognitive-behavioural treatment for rapists: Can we do better?' (1999) 19 ClinPsychol Rev 875.

often exhibit significantly reduced rates of re-offending. It must be acknowledged that the evidence supporting this may be due partially to the fact that there are more examples of CBT in action,⁴⁹ and with further experimentation or examining the validity of previous results, it may become apparent that other techniques are equally or even more effective. However, at present CBT interlinked with RNR are the most effective sex offender treatment programs in terms of reducing the re-offending of sex offenders in general.

Although the evidence demonstrates RNR based CBT are the most effective treatments for sex offenders, when examining the subcategories of sex offenders it becomes apparent that different sex offenders require emphasis on different aspects of these programs. Measuring the effectiveness of sex offender treatment programs for different types of sexual offenders is not restricted to rates of re-offending,⁵⁰ in fact, identifying differences in dynamic risk factors such as the levels of motivation of offenders before and after treatment, often gives more insight into where studies are effective and which aspects require development. By first identifying the risk factors, it is simpler to determine for each subcategory which areas require the most focus, in other words following the concept of the need principle. A study by Barrett et al assessed changes in motivation levels before and after treatment for child molesters, paedophilic child molesters and those who committed offences against adults, namely rapists.⁵¹ Non-Pedophilic child molesters had consistently higher motivation levels following treatment compared to initial analysis, whereas paedophilic child molesters and rapists displayed more complex results. Following treatment, admission of guilt and acceptance of responsibility for both categories were substantially lower than expected. Before identifying possible ways to increase

⁴⁹ Schmucker&Lösel (n 35).

⁵⁰ Kurt M Bumby, 'Assessing the Cognitive Distortions of Child Molesters and Rapists: Development and Validation of the MOLEST and RAPE Scales' (1996) 8 Sex Abuse 37.

⁵¹ Barrett (n 9) 279.

results in these areas the reasons for both categories' results and the theories behind them should be investigated.

Though both paedophilic child molesters and rapists are often diagnosed with sexual disorders, evidence has shown stronger sexual deviance among paedophilic child molesters.⁵² What can be extracted from this knowledge is that paedophilic child molesters are more likely to have cognitive distortions about sexual deviance, that is to say, have beliefs about sex that are not socially acceptable and can lead to criminal behaviour.⁵³ CBT examples that focus on these aspects, such as The Kia Marama Sex Offender Treatment Program, place extra emphasis on tackling sexual deviance by not only cognitive restructuring but largely through introducing coping methods, which do not attempt to dispel distortions, but allow paedophilic child molesters to develop ways to resist deviance.⁵⁴ Furthermore due to the evidence linking paedophilic child molesters with sexual disorders, it is perhaps wise to consider also increasing levels of behavioural treatments such as masturbatory reconditioning.⁵⁵ These treatments illustrate the importance of identifying the need factor of the RNR theory within CBT, in an attempt to change the offender's perception of what is arousing from sexually deviant to more socially acceptable concepts, such as consenting sex between adults.⁵⁶

While rapists are often viewed as having sexual disorders,⁵⁷ studies have shown that they do not benefit solely from treatments aimed at sexually deviant behaviour and often do not exhibit particularly successful outcomes when

⁵² Donaldo Canales, Mark Olver & Stephen Wong, 'Construct Validity of the Violence Risk Scale-Sexual Offender Version for Measuring Sexual Deviance' (2009) 21 Sex Abuse 474.

⁵³ Bumby (n 50) 38.

⁵⁴ Stephen Hudson, David Wales, Leon Bakker & Tony Ward, 'Dynamic Risk Factors: The Kia Marama Evaluation' (2002) 14 Sex Abuse 103.

⁵⁵ Barrett (n 9) 281.

⁵⁶ Howitt (n 1).

⁵⁷ Barrett (n 9) 280.

treated alongside child molesters in the SOTP.^{58,59} In fact, they are possibly the most complex and intriguing sector of sex offenders, in regards to sex offender treatment programs effectiveness. Eher et al noted after reviewing diagnoses of rapists, that 76% had personality disorders rather than sexual disorders, with particularly high rates in antisocial disorders.⁶⁰ Some academics are of the opinion that sex offender treatment programs among those with personality disorders are not only ineffective, but can actually have harmful effects. Jones, who describes these outcomes as iatrogenic, argues that sex offenders with personality disorders have a significantly high dropout rate (50%) and even when completed, sex offender treatment programs do not result in a significant reduction in re-offending rates.⁶¹ He continues to voice his concerns that if treatments are ineffective the first time they are experienced, cognitive distortions could be strengthened and a lack of belief in the ability to change could be instilled in the offender. Evidence to the contrary shows that there is statistical uncertainty and not enough evidence exists at present to support this theory.⁶²

Regardless, when treating offenders with personality disorders (in this example rapists), different methods should be applied to increase effectiveness to its fullest potential. In response to antisocial disorders, methods should heavily rely on measures to decrease anti-social behaviour⁶³ such as anger control and interpersonal problem solving.⁶⁴ There is also evidence illustrating that when treating rapists it is prudent to consider Lang's triple response theory,⁶⁵ this advises that

⁵⁸ STEP Team, 'The Therapeutic Impact of Sex Offender Treatment Programmes' (1995) 42 *Probat* J 2.

⁵⁹ Sophie Davison & Aleksandar Janca, 'Personality Disorder and Criminal Behaviour. What Is the Nature of the Relationship?' (2012) 25 *Curr Opin Psychiatr* 39, 41.

⁶⁰ Eher et al, (2010) cited in Davison & Janca (n 59).

⁶¹ Jones (n 8) 72.

⁶² D'Silva (n 27).

⁶³ Barrett (n 9) 280.

⁶⁴ Lipsey (n 42).

⁶⁵ (1968) cited in Marx (n 48) 882.

there are “three response channels (verbal, overt motor and physiologic)”⁶⁶ and that these channels should be measured and evaluated in regards to what motives different rapists do offend. Applying his theory in practice is important when evaluating which motivation function is primary in different offenders. For example, an individual may find a victim in pain sexually arousing at all times but this arousal alone may not lead to offending.⁶⁷ However, combining this sexual deviance with certain conditions, such as high-tension levels, may lead to undesirable acts being carried out.⁶⁸ This theory would imply that significant inclusion of techniques to encourage offenders to accept that, although they may have deviant thoughts, they can learn coping methods so as not to act on them, would perhaps prove more effective.⁶⁹

This is further illustrated in a study by Marshall and Barbaree in which the results from 60 rapists showed that 30% were equally or more aroused by scenes of nonconsensual sex than consensual sex.⁷⁰ When compared with 27% of non-offenders finding nonconsensual scenes equally or more arousing than consensual sex, the difference is insignificant. This further demonstrates that rapists are not necessarily more likely to have sexually deviant thoughts, and that sex offender treatment programs should also focus on other factors to increase effectiveness. Sexual deviance clearly does not play a primary role in motivation of rapists and therefore it would be less effective to rely purely on behavioural treatments. It is evident that despite CBT displaying significant reductions in rates of re-offending, the term and concepts of ‘cognitive-behavioural’ programs are not specific. This runs concurrently with the view that sex offenders do not commit offences for the same reasons and therefore treatments that are generalised are very unlikely to

⁶⁶ Marx (n 48) 882.

⁶⁷ *ibid.*

⁶⁸ *ibid.*

⁶⁹ *ibid.*

⁷⁰ William L. Marshall & James McGuire, ‘Effect sizes in treatment of sexual offenders’ (2003) 47*Int J Offender Ther Comp Criminol* 653.

be effective for all categories.⁷¹ When assessing sex offender treatment programs tailored to support different subcategories of sex offenders, it is important to understand theories of why different types of sex offenders, particularly rapists, offend and apply these to sex offender treatment programs in practice in order to induce maximum effectiveness. This again, illustrates the importance of identifying the need factor of individual offenders, as suggested in the RNR theory.

In conclusion, the importance of identifying the factors of risk, need and responsivity is evident in selecting treatment programs that are most likely to be effective for sex offenders in general. The principles of the RNR theory allow for treatment programs that are specifically tailored to the diverse needs of individual sex offenders. These principles, when interlinked with a combination of both cognitive and behavioural treatments, are decidedly the most effective in practice and this is supported by numerous studies, which have concluded that RNR based CBT significantly reduces re-offending for sex offenders.⁷² When more closely assessing the effectiveness of sex offender treatment programs, it is also important to note that the reasons behind offending differ hugely and sex offenders cannot be effectively treated identically. Following the RNR model allows for specific treatment of subcategories that at first appear similar, but which require largely different methods of treatment to achieve positive results. When applying the RNR principles to CBT, hugely beneficial treatment of sex offenders can occur, these treatment programs can again provide specific treatments based on the most effective ways to correct specific motivations behind offending, either cognitive or behavioural.⁷³ It is clear that extensive research into theories behind motivations and possible ways of reducing re-offending rates, particularly among rapists, is needed, and

⁷¹ STEP Team (n 58).

⁷² Howitt (n 1).

⁷³ STEP Team (n 58).

present evidence of success is by no means conclusive. However when correctly applying the RNR theory to CBT, sex offender treatment programs can not only provide hugely desirable intrinsic outcomes for offenders, but also significantly diminish the likelihood of re-offending in the community.⁷⁴

⁷⁴ Andrews (n 2).

A Critique of Enlightened Shareholders Value: Revisiting the Shareholders' Primacy Theory

Collins C. Ajibo*

Abstract

The statutory re-conceptualisation of the traditional common law shareholders primacy into 'enlightened shareholders value' emblematic of section 172 of the UK Company Act 2006 has generated a universe of views among scholars. While some scholars hypothesise that the enlightened shareholder value concept epitomised by section 172 is no more than a re-affirmation of the traditional common law shareholders' primacy under different guise, others theorise that it inches towards the 'pluralist theory' of continental European tradition. Still further, others argue that the re-conceptualisation is emblematic of the convergence of principles from both the traditional 'primacy theory' and the 'pluralist theory'. While the preceding divergences on the contours of the substantive rules subsist, arguments and counter-arguments have similarly underpinned the potential fallout of the procedural framework for the enforcement of the section consequent on procedural constraints. It is contended that despite the laudable statutory re-conceptualisation, section 172 only added little improvement (if any) on the traditional common law shareholders' primacy; and its greatest shortcoming lies on the enforcement constraints. Nevertheless, the courts can still adopt a teleological interpretative approach to the construction of the section that plug the loophole in the stakeholders' protection.

I. Introduction

The graduation from the traditional common law shareholders' primacy to enlightened shareholders' principle now encapsulated in section 172 of the UK Company Act 2006 has led to so many postulations as to the potential implication of the section.¹ The Company Law Review

* LLB, LLM Manchester (Distinction), PhD Candidate, School of Law, University of Manchester. Email: ajibochikodili@yahoo.com; chikodili.ajibo@postgrad.manchester.ac.uk. Note that this article (with

Steering Group(CLRSG)²considered the pluralist approach characteristic of some continental European countries³ but abandoned it in favour of enlightened shareholders approach. The directors are now constrained to have regard to stakeholders' interests as well as to observe the tenets of corporate social responsibilities while promoting the success of the company for the benefit of members as a whole. However, the preceding responsibilities have equally thrown up conceptual arguments of various strands. Thus, while some scholars postulate that the enlightened shareholder concept epitomised by section 172 is coterminous with traditional shareholders' 'primacy theory' that underpinned the common law paradigm,⁴ others hypothesise that it inclines towards 'pluralist theory' characteristic of a significant number of continental European countries.⁵ Still further, others argue that the re-conceptualisation is emblematic of a confluence of principles from both the traditional 'primacy

slight modifications) was first published in *Birkbeck Law Review*, April Issue, 2014.

¹ Andrew Keay, 'The Enlightened Shareholder Value and Corporate Governance', [2005], *Modern Law Review*, 76(5), 940. On the relevance of shareholders' primacy in Anglo-American corporate law and the need for a stakeholders' theory of corporate governance see, Jingchen Zhao, 'The Curious Case of Shareholder Primacy Norm: Calling for a more Realistic Theory', [2012] 15 *ITBLR* 2.

² The Government Explanatory Notes to Company Act 2006 indicates that section 172 is a product of effort to codify the directors' duties in statutory form, unlike the common law position that was largely not codified. See The UK Government, 'Explanatory Notes to Company Act 2006' <<http://www.legislation.gov.uk/ukpga/2006/46/notes>> accessed: 11 February 2014.

³ Continental European countries with pluralist approach include, inter alia, Germany, Netherland, Austria, among others. See Luca Cerioni, 'The Success of the Company in S.172 (1) of UK Company Act 2006: Towards an 'Enlightened Directors' Primacy?' (2008) *Original Law Review* 4, 37.

⁴ David Collison *et al*, 'Financialization and Company Law: A Study of the UK Company Law Review' [2013] *Critical Perspectives on Accounting*, <<http://www.sciencedirect.com/science/article/pii/S1045235412001311#>> accessed: 25 November 2013.

⁵ Roman Tomasic, 'Company Law Modernisation and Corporate Governance in the UK - Some Recent Issues and Debates' [2011] *Victoria Law School Journal*, 43; Luca Cerioni, 'The Success of the Company in S.172 (1) of UK Company Act 2006: Towards an 'Enlightened Directors' Primacy?' [2008] *OLR* 4.

theory' and 'pluralist theory' typical of continental Europe.⁶ While the arguments on the contours of the substantive rules linger, others have gone further to express scepticism on the enforcement of shareholders' and/or stakeholders' rights consequent on procedural constraints. These divergences tend to challenge the philosophical premise behind the reformulation of the section; and may, as well, have far reaching implications on the practical fallout of its construction. A critique of the foregoing trajectories therefore forms the focus of this article. It will be argued that despite the laudable statutory re-conceptualisation, section 172 only added little improvement (if any) on the shareholder 'primacy theory'; and its greatest shortcoming lies on the procedural enforcement constraints. However, the courts can salvage the situation through teleological interpretation of the section.

The article is divided into eight sections to underscore the significance of each of the concepts as well as to enhance better understanding by the readers. Section II examines the conceptual premise of shareholders versus stakeholders' debate. Section III lays out the constituents of section 172 for easy reference, then examines the background preceding the reform before analysing the contours of 'promoting the success of the company' in relation to the common law position. Section IV dwells on the concept of good faith. Section V then examines the core concept of enlightened shareholders as it relates to the conjunctive terms of 'have regard (amongst other things) to'. Section VI explores whether the position of stakeholders particularly employees and creditors' is better protected under the current framework compared to traditional common law position. Section VII then argues that the controversy characterising the section could be ameliorated if the courts could adopt a teleological interpretative paradigm. Section VIII concludes

⁶ Stephen Brammer *et al.*, 'Corporate Social Responsibility and Institutional Theory: New Perspectives on Private Governance' [2010] *Socio-Economic Review* 10 (1), 12.

that section 172 added little (if anything) to the traditional common law shareholder primacy theory, except if the courts could adopt a purposeful interpretative approach anchored on a teleological paradigm.

II. Conceptual Framework of Shareholders versus Stakeholders Debate

Before delving into the trajectories of section 172, it would be pertinent at the outset to examine in general the competing theories of corporation to underscore the normative foundation cum divergent postulations that underpin the shareholder versus stakeholder debate. The doctrinal conception of corporation in early years had been an artificial entity (a derivative of concession or grant theory of corporation)⁷ which came into being by virtue of substantive laws of the state – demonstrated in chartered nature of corporation then.⁸ However, as the state continued to loosen legal strictures making it easier for the incorporation of companies, a new doctrinal conception emerged therefrom, in the form of natural entity theory.⁹ In contrast to an artificial entity, natural entity theory postulates that a corporation is a natural creation consequent upon agglomeration of private initiatives of the individuals, and thus can only exercise such powers as are extended to it by the shareholders. The state should refrain from imposition of regulatory constraints as the corporation, governed by private law, exists to serve shareholders' interests as against stakeholders' interests.¹⁰

⁷ Morton J Horwitz, 'Santa Clara Revisited: The Development of Corporate Theory' [1985-1986] *West Virginia Law Review* 88, 184.

⁸ David Millon, 'Theories of Corporation' [1990] (2) *Duke Law Journal*, 206; David Millon, 'New Directions in Corporate Law: Communitarians, Contractarians, and the Crisis in Corporate Law' [1993] 50 *Washington & Lee Law Review* 1373.

⁹ Millon, 'Theories of Corporation' [1990] *Duke Law Journal* 201.

¹⁰ For a review of these theories see, David Millon, 'Theories of Corporation', 211-216. Note that aggregate theory of corporation provided the theoretical basis for natural entity theory but was unsustainable due to the separation of ownership from management as well as the crystallisation

This doctrinal premise provided the theoretical framework for the shareholder versus stakeholder debate between Berle¹¹ and Dodd¹² in the early 1930s that significantly shaped subsequent legal discourse on the scope of directors' duties. Neo-classical proponents of Berle's position contend that shareholder value maximisation constitutes the only theory of corporation consistent with free market economy.¹³ In their provocative article, *The End of History for Corporate Law*,¹⁴ Hansmann and Kraakman, in supporting the shareholders' primacy theory, posit that other stakeholders such as creditors, employees, customers, suppliers, environments, among others, can only be part of the corporate governance equation if they are party to the express and unambiguous contract with the corporation. Failing that, they can only lay claim to protections of other bodies of law, otherwise their interests are not to be the concern of corporate management. Specifically, corporate accountability of directors is owed to shareholders only who invested their money in the corporation and not otherwise. Thus, members of society affected by corporate activities may avail themselves of the protection of, for example, environmental law or the law of tort, but they are not to expect corporate management to have them in contemplation during the operation of the company.¹⁵

In contrast to the above, progressive writers, particularly Cynthia William, argue that in the context of increasing globalisation, the argument that corporation exists purely to

of majority rule as a voting procedure for shareholders as against the prior unanimity rule.

¹¹ A Berle, 'For Whom Corporate Managers Are Trustees: Note' [1932] 45(8) Harvard Law Review 1365.

¹² E Merrick Dodd, 'For Whom Are Corporate Managers Trustees?' [1932] 45(7) Harvard Law Review 1145.

¹³ Stephen Bainbridge, 'In Defence of Shareholder Wealth Maximization Norm' [1993] 50 Washington & Lee Law Review 1423; Milton Friedman, 'The Social Responsibility of Business is to Increase its Profit' *New York Time Magazine* (1970).

¹⁴ Henry Hansmann and Reinier Kraakman, 'The End of History for Corporate Law' [2001] 89 *Georgetown Law Journal* 442.

¹⁵ *ibid* 442.

maximise the shareholders' wealth is illusory.¹⁶ In other words, the so-called regulatory rules and explicit contract to regulate company relationship with stakeholders are inadequate. Equally, Margaret Blair and Leon Stout question the validity of shareholder primacy theory.¹⁷ Blair and Stout argue that the existence of corporation is a by-product of team production involving the input of various interests not limited to shareholders *per se*.¹⁸ In fact, the team production theory rejects shareholder primacy theory.

Shareholder versus stakeholder debates can further be analysed from the standpoint of the contractarianism versus communitarianism debate;¹⁹ the predominant versus the progressive position;²⁰ the monism versus pluralism debate, among others, reflecting the diversity of perspectives.²¹ Thus, these conceptual frameworks shaped the evolution and continue to shape the growth of corporate law and by implication the directors' duties which are the touchstone of corporate governance. It can be argued therefore that the statutory re-conceptualisation of section 172 Company Act 2006 might have been influenced in part, at least conceptually, by the preceding. Analysis of the core substantive import of the section forms the focus of the following parts punctuated, first and foremost, by laying out the constituents of the section for easy reference.

¹⁶ Cynthia William, 'Corporate Social Responsibility in Era of Globalisation' [2001-2002] 35 *University of California Davies Law Review*, 720.

¹⁷ Margaret Blair and Lynn Stout, 'Team Production Theory of Corporate Law' [1999] 85 *Virginia Law Review* 246, 278.

¹⁸ Blair and Stout, 'Team Production Theory of Corporate Law', 280.

¹⁹ Benedict Sheehy, 'Scrooge the Reluctant Stakeholder: Theoretical Problems in the Shareholder-Stakeholder Debate' [2005-2006] 14 *University of Miami Business Law Review*, 226.

²⁰ William, 'Corporate Social Responsibility in Era of Globalisation' 711-16.

²¹ Thomas W Dunfree, 'Corporate Governance in a Market with Morality' [1999] 62 *Law Contemporary Problems*, 130.

III. Section 172

Section 172 provides:

(1) A director of a company must act in the way he considers, in good faith, would be most likely to promote the success of the company for the benefit of its members as a whole, and in doing so have regard (amongst other matters) to -

- a) the likely consequences of any decision in the long term,
- b) the interests of the company's employees,
- c) the need to foster the company's business relationships with suppliers, customers and others,
- d) the impact of the company's operations on the community and the environment,
- e) the desirability of the company maintaining a reputation for high standards of business conduct, and
- f) the need to act fairly as between members of the company.

(2) Where or to the extent that the purposes of the company consist of or include purposes other than the benefit of its members, subsection (1) has effect as if the reference to promoting the success of the company for the benefit of its members were to achieving those purposes.

(3) The duty imposed by this section has effect subject to any enactment or rule of law requiring directors, in certain circumstances, to consider or act in the interests of creditors of the company.

IV. Background to the Company Law Reform of 2006

The company law reform of 2006 was an agglomeration of the efforts to reform the UK corporate governance model, inspired collectively by the Cadbury Committee Reports of 1992; the Greenbury Committee Report of 1996; and the Hampel Committee Report of 1998, which consolidated the preceding into a Combined Code.²² Following the preceding Reports,²³ the Department of Trade and Industry (DTI)²⁴ commissioned Company Law Review Committee charged with modernising the UK company law to make it “simple, efficient and cost effective framework for British business in the twenty-first century”.²⁵ The Committee submitted its Final Report involving the recommendation options to the Secretary of State for Trade and Industry on 26 July 2001. After series of further consultations, the government finally issued the White Papers “Modernising Company Law” (July 2002) and “Company Law Reform” (March 2005).²⁶ Following further public comments on the Government intention as embodied in the White Paper of 2005, the Government finally introduced the Company Law Reform Bill to the House of Lords on 4 November 2005, setting the stage for the reform that included section 172 which constitutes the focal point of this article. The analysis of the substantive parts of the section forms the focus of next discussion hereunder.

²² Note also that other Reports such as those emanating from the Turnbull Committee, and the Higgs Review, amongst others, added to the build up to the Company law reform agendas. These Reports in one way or the other tended to offer recommendations to improve the UK corporate governance model including the workings of the board of directors and disclosure of directors’ remunerations. Note equally that the Combined Code on Corporate Governance has been replaced by the UK Corporate Governance Code of June 2010. See Derek French et al, *Company Law*, (Oxford University Press, Oxford, 2012-2013), 429-430

²³ Note that these Reports were not limited to the reform of section 172 *inter se*, but straddle the principles of corporate governance in its entirety.

²⁴ Note that the DTI has now been replaced by the Department of Business, Enterprise and Regulatory Reform and the Department of Innovation, Universities and Skills in 2007.

²⁵ See The UK Government, ‘Explanatory Notes to Company Act 2006’.

²⁶ See The UK Government, ‘Explanatory Notes to Company Act 2006’.

V. Promoting the Success of the Company

Traditionally, directors under common law owe the fiduciary duty to act bona fide for the interest of the company.²⁷ This, by and large, translates into balancing the short term interests of present members with the long term interests of future members²⁸ with shareholders' interest construed from the prism of advancing shareholders' value.²⁹ The statutory reformulation substituted 'interest of the company' with 'promoting the success of the company'. However the Act offers no definition of what constitutes promoting the success of the company for the benefit of members as a whole. This is likely to create a significant problem for the court in construing it.³⁰

Section 170(4) provides that the general duties shall be interpreted and applied in the same manner as common law rules and equitable principles, and regard shall be had to the corresponding common law rules and equitable principles in interpreting and applying the general duties.³¹ This could mean that the meaning ascribed to 'interest of the company' under the common law namely –the benefit for the present and future shareholders³² may be relevant in interpreting 'promoting success of the company for the benefit of its members as a whole', embedded in section 172.

Although parliamentary debate³³ described success as long term increase in shareholders' value evidencing economic

²⁷ *Percival v Wright* [1902] 2 Ch 421.

²⁸ Brenda M Hannigan, *Company Law* (2nd edn Oxford University Press, Oxford, 2009) 205.

²⁹ See *Smith & Fawcett Ltd* [1942] Ch 304, 306; *Greenhalgh v Aderne Cinemas Ltd* [1951] Ch286.

³⁰ See *R (on the application of People & Planet) v HM Treasury* [2009] EWHC 3020 (Admin).

³¹ See Companies Act 2006, sub-s 170(4).

³² It has been noted that contrary to argument postulated elsewhere, interest of the company under common law was not limited to short term interest but include also long term interest. See Paul L. Davies, *Gower and Davies Principles of Modern Company Law* (8th edn Sweet and Maxwell, London, 2008) 510.

³³ *Pepper v Hart* [1993] AC 593.

success of the company,³⁴ that by no means constitutes a clear-cut parameter of measurement. It has been posited that promoting the success of the company for the benefit of its members equates to observance by the directors of the objectives of the company set out in its constitution.³⁵

However, where a company is a charity and/or community interest-based,³⁶ or incorporated for a specific object, determining the success of the company may be difficult. Although it has been suggested that the measurement of success of the company in such circumstances could be determined based on achieving the intended object,³⁷ that by no means settles the issue, particularly where the 'intended object' does not significantly incorporate external constituencies. Indeed, such non-commercial or charitable companies might, perhaps, be required to promote the success of the companies to achieve the charitable objects having regard to stakeholders' interest. Nonetheless, success would still be difficult to calibrate where companies exist for interest other than that of its members.³⁸ As noted,³⁹ subsection 2 is intended to preserve non-commercial objectives of companies as illustrated by *Horsley v Weight*.⁴⁰ It remains to be seen how courts will construe this subsection given "different hierarchy of priorities"⁴¹ that have to be weighed in determining the success of the company.

One potential guide to determining the success of the company might be that directors' actions must not *ipso facto*

³⁴ See HL Deb. Vol 678 GC 255-8(6 Feb 2006).

³⁵ See Hannigan, *Company Law* at 209.

³⁶ See DTI, Companies Act 2006: Duties of Directors: Ministerial Statement (DTI, London, June 2007) 7, statement of Lord Goldsmith, Lords Grand Committee, 6 Feb Column 255. See also, s172(2) UK Company Act 2006 (hereinafter referred to as CA 2006).

³⁷ Hannigan, *Company Law* at 209.

³⁸ Hannigan, *Company Law* at 209.

³⁹ A Alcock, 'Accidental Change to Directors' Duties?' [2009] 30(12) *Company Lawyer*, 7.

⁴⁰ *Horsley v Weight* [1984] 3 All ER 1045.

⁴¹ Paul Omar, 'In the Wake of Company Act 2006: An Assessment of Potential Impact of Reforms to Company Law' [2009] 20(2) *International Company and Commercial Law Review*, 44.

promote the success of the company in objective terms as long as they acted in good faith believing their action to be likely to promote the success of the company.⁴² This argument is, perhaps, reinforced by the Guidance on Key Clauses in Company Law Reform Bill which states that “[t]he decision as to what will promote success and what constitutes such, is one for the directors’ good faith judgment...”.⁴³ The foregoing no less encapsulates wide discretionary powers of the directors in the exercise of the management judgment.

Nonetheless, where directors’ action is devoid of consideration of company’s interest and there is no ground on which he or she could reasonably arrive at such a conclusion that it was done for company’s interest, he or she would be in breach of duty.⁴⁴ Furthermore, directors’ action would not be justified where, irrespective of the fact that the boards would have arrived at the same decision as the directors, their action, considered in totality, could not be said to be in the interest of the company.⁴⁵ This however does not detract from the wide discretion of directors embedded in section 172(1) which must be exercised in good faith.

It is settled that embedded in section 172(1) is a subjective element⁴⁶ in promoting the success of the company for the benefit of the whole members in contrast to the common law duty that directors “act in good faith for the benefit of the company which has element of objectivity”.⁴⁷ A subjective test may prove cumbersome for the court to construe. Consequently, it has been postulated that just like under

⁴² Andrew Keay, ‘S.172 (1) of the Companies Act 2006: an Interpretation and Assessment’ [2007] 28(4) *Company Lawyer*, 5.

⁴³ See DTI, Guidance on Key Clauses to the Company Law for a Competitive Economy: Developing the framework (DTI, London 2000) para 63.

⁴⁴ See *Item Software UK Ltd v Fassihi* [2005] 2 BCLC 91.

⁴⁵ *W & W Roith Ltd Re* [1967] 1 WLR 432.

⁴⁶ Davies, Gower and Davies Principles of Modern Company Law at 510.

⁴⁷ Ji L. Yap, ‘Considering the Enlightened Shareholders Value Principle’ [2010] 31(2) *Company Lawyer*, 3.

common law, courts will ultimately introduce an objective element in construing acting in ‘good faith’ to promote the success of the company for the benefit of the members as a whole.⁴⁸ Thus, illustrative of this position was the case of *Chatterbridge Corp. Ltd. V Lloyd Bank Ltd.*,⁴⁹ where it was held, inter alia, that the duty to act in good faith for the company’s interest could be faulted where the actions of the directors could not be reasonably considered to be in the interest of the company by any reasonable and intelligent person.

Apart from the stakeholders’ interest enumerated in subsection 1 to section 172, directors should, in promoting the success of the company, have regard to the need to act fairly between members of the company.⁵⁰ Subsection 1(f) in effect probably entails that directors have to take into consideration the effect of their proposal on different classes of shareholders,⁵¹ and act fairly among all in a manner absent protecting sectional interest.⁵² Failure of the directors in this regard may trigger minority action under unfair prejudice provision. However, it would appear that directors could be justified under section 172(1) to defend any decision to promote the success of the company for the benefit of the whole members even if it favours employees’ interest over short term profit.⁵³

Similarly, where directors’ actions will ultimately promote the success of the company but, invariably, affect certain shareholders, their action could still be upheld notwithstanding. Thus, in *Mutual Life Insurance Co. of New*

⁴⁸ Keay, ‘S.172 (1) of the Companies Act 2006: an Interpretation and Assessment’, 5.

⁴⁹ *Chatterbridge Corp Ltd v Lloyd Bank Ltd*[1970] Ch 62.

⁵⁰ See subsection 1(f). See also *R. (on the application of People & Planet) v HM Treasury* [2009] EWHC 3020 (Admin).

⁵¹ See *Re BSB Holding Ltd. No2* [1996] 1 BCLC 155.

⁵² See *Mills v Mills* [1930] 60 CLR 150.

⁵³ John Lowry, ‘The Duty of Loyalty of Company Directors: Bridging the Accountability Gap through Efficient Disclosure’ [2009] 68(3) *Cambridge Law Journal*, 607; see also *Re Wellab Engineers Ltd* [1990] BCLC 833 per Hoffmann J.

York v Rank Organisation Ltd.,⁵⁴ despite discrimination in issue of shares to hedge cost of regulatory compliance by the company, it was held, inter alia, that directors honestly believed that raising capital as such would benefit the company with consequential benefit to all the shareholders.

It should be noted that the duty to promote the success of the company for the benefit of the members as a whole is subject to subsection 3; and as argued, some otheroverriding legislations particularly employment, consumer, safety and discrimination legislations, even if non-compliance with the above legislations will promote the success of the company for the benefit of members as whole.⁵⁵

VI. Good Faith

Good faith is not defined by the Act. Thus, since regard must be had to common law rules and equitable principles in construing the section⁵⁶ it would appear that the directors' discretion embodied in the section, in similar fashion as under common law, is uninhibited. As illustrated by Lord Greene MR in *Smith & Fawcett Ltd.*,⁵⁷ directors are bound to exercise their powers "bona fide in what they consider, not what a court may consider, is in the interest of the company". Although the preceding ruling might be informed by the fact that directors are better positioned to make value judgments on what course the company might take, such an unfettered discretion might equally lead to potential abuse of powers by the directors.⁵⁸

⁵⁴ *Mutual Life Insurance Co. of New York v Rank Organisation Ltd*[1985] BCLC 11.

⁵⁵ Davies, Gower and Davies *Principles of Modern Company Law* at 510; see also, s.172 (3) CA 2006.

⁵⁶ See s.170 (4) CA 2006.

⁵⁷ [1942] Ch 304, at 306 (CA).

⁵⁸ See *Wrexham Associated Football Club Ltd (In Administration) v Crucialmove Ltd*[2008] 1 B.C.L.C. 508, where the director, in exercise of discretion, was involved in a breach of fiduciary duty as well as conflict of interest as result of personal gain.

Guidance on Key Clauses in Company Law Reform Bill⁵⁹ stipulates that good faith should be exercised in a manner characteristic of a reasonable man of skill, care and diligence. Nonetheless, it has been postulated that directors ought not to be liable in breach of their fiduciary duties where their actions are unreasonable but he or she honestly believes it is done in good faith for the interest of the company.⁶⁰ This is because good faith is, inter alia, anchored on honesty and loyalty, not necessarily competence.⁶¹ However, courts may invalidate directors' actions done for collateral purpose,⁶² irrespective of the powers of the boards to ratify subsequently.

Moreover, as argued, where a director fails to exercise power in good faith to promote the success of the company, their action may be open to review; and if loss resulted to the company, he or she will be liable to make good the loss.⁶³ In this regard, non-executive directors have been advised not to allow themselves to be dominated by directors lest they would not be able to convince the court that they acted in good faith.⁶⁴ It can be argued, on the contrary, that where bad faith characterises action of the directors, it ought to ground liability.

VII. Enlightened Shareholder Principle

The common law position was based on primacy of shareholders value⁶⁵ - what can be termed the 'primacy theory'. It is arguable whether section 172 (1) encapsulated as

⁵⁹ See DTI, Guidance on Key Clauses to the Company Law for a Competitive Economy: Developing the Framework at 63.

⁶⁰ DKC Wu, 'Managerial Behaviour, Company Law and the Problem of Enlightened Shareholders Value' [2010] *Company Lawyer*, 2.

⁶¹ *ibid* 2.

⁶² See *Ultraframe (UK) v Fielding* [2005] EWHC 1638 Ch.

⁶³ John Birds *et al*, *Company Law* (7th edn Jordan Publishing, Bristol, 2009) 597.

⁶⁴ Ji L Yap, 'Hear no Evil, See no Evil, Speak no evil: the Total Inactivity of Non-Executive Directors' [2009] 20(11) *International Company and Commercial Law Review*, 412.

⁶⁵ *Greenhalgh v Aderne Cinemas Ltd* [1951] Ch286; see also *Hutton v West Cork Rly Co* [1883] 23 ChD 654.

‘enlightened shareholders value’ significantly differed from the existing legal position. According to Davies,⁶⁶ section 172 is an improvement on common law but only a modest one. However, it has been held that the section did “little more than set out the pre-existing law”.⁶⁷ Although this decision has been challenged as not reflecting the correct legal position,⁶⁸ it does indicate the divergence that underpins this sphere. Nonetheless, it has been contended that section 172(1) is in accord with OECD Principles on Corporate Governance which emphasises, among other things, cooperation between the corporation and stakeholders in creating wealth, and that the full import of such accord will come into effect upon directors’ internalisation of the section.⁶⁹

It is considered that section 172 (1) represents a significant improvement, since amongst other things, the interest of a wider scope of stakeholders will now be considered by directors than hitherto the case. However, it would be hard to claim that the section replicates the ‘dual consideration theory’ in its entirety characteristic of significant number of continental European countries, particularly the German co-determination model that accords dual consideration to both shareholders and stakeholders in management decisions.

Undoubtedly, the directors in promoting the success of the company for the benefit of the members as a whole should have regard to, the likely consequences of any decision in the long term;⁷⁰ interest of the employees;⁷¹ interest

⁶⁶ See Davies, Gower and Davies Principles of Modern Company Law at 509.

⁶⁷ See *Re West Coast Capital Ltd* [2008] CSOH 72 per Lord Glennie; he reasoned that sections 171 - 172 did little more than set out the pre-existing legal position.

⁶⁸ Robert Goddard, ‘Directors Duties’ [2008] 12 *Edinburgh Law Review*, 468; Lowry, ‘The Duty of Loyalty of Company Directors: Bridging the Accountability Gap through Efficient Disclosure’ [2009] 68(3) *The Cambridge Law Journal* 607, 614 (arguing that s172 is a step further than pre-existing law).

⁶⁹ Cerioni, ‘The Success of the Company in S.172 (1) of UK Company Act 2006: Towards an ‘Enlightened Directors’ Primacy?’ [2008] OLR 4(1)

⁷⁰ Companies Act 2006, s 172(1)(a).

of suppliers, customers and others;⁷² impact of their operation on the environment; the need for high standards of conduct;⁷³ and the need to act fairly between members of the company.⁷⁴ It is settled that the stakeholders' interest encapsulated in the above subsection by no means equates to shareholders' interest.⁷⁵ Rather, directors are called upon to take into account stakeholders' interests so long as such action will promote the success of the company for the benefit of members as a whole.

Under common law, it was possible for directors to take into account stakeholders' interests so long as it promoted the interest of the company for the benefit of shareholders as a whole.⁷⁶ This was illustrated by the case of *Hutton v West Cork Rly*,⁷⁷ where it was held that, "the law does not say that there are to be no cakes and ale, but there are to be no cakes and ale except such as are required for the benefit of the company".

The stakeholders' interest may appear, prima facie, to be better protected under section 172 than hitherto the case. However, the shortcoming of section 172 (1) becomes more apparent when one inquires as to whether stakeholders can directly enforce the observance of their interest. Consequently, it has been postulated that except shareholders, or where a stakeholder doubles as a shareholder, other stakeholders lack the capacity to enforce the observance of their interests embodied in the

⁷¹ Companies Act 2006, sub-s 172(1)(b).

⁷² Companies Act 2006, sub-s 172(1)(c).

⁷³ Companies Act 2006, sub-ss 172(1)(d - e).

⁷⁴ Companies Act 2006, sub-s 172(1)(f).

⁷⁵ See Davies, Gower and Davies Principles of Modern Company Law at 510.

⁷⁶ See the case of *Evans v Brunner, Mond and Company, Limited* [1921] 1 Ch 359, where the company expenditure for scientific research incidental to its main object of chemical manufacturing was held to be within the object of the company and therefore not ultra vires.

⁷⁷ *Hutton v West Cork Rly* [1883] 23 Ch D 673, per Bowen LJ.

section.⁷⁸ Similarly, shareholders bringing derivative action on company's behalf must obtain court's approval,⁷⁹ as well as encounter other (almost insurmountable) hurdles.⁸⁰ The above shortcomings have led to questions as to whether the lots of shareholder and stakeholders are better protected under section 172 than hitherto the case. It has been argued that the inability of the stakeholders to enforce the observance of the section 172 (1) directly may entail that the section will hardly be litigated.⁸¹ It could be argued that, perhaps, empowering direct stakeholders' enforcement could lead to vexatious actions against directors. Nonetheless, a duty is only useful in law if it is enforceable.⁸²

The encapsulation of what is generally known as corporate social responsibility (CSR) in section 172 (1) (d - e) is a giant stride in law making. However, as noted, corporate governance that is shareholder-centric as is the case with section 172 (1) may not adequately cater to CSR.⁸³ Nonetheless, it has been argued that shareholders are deemed "to be enlightened and to want to take into account issues of [CSR]"⁸⁴

⁷⁸ Companies Act 2006, s 261; Parker Hood, 'Directors Duties under the Companies Act 2006: Clarity or Confusion?' [2013] 13 *Journal of Corporate Law Studies*, 47.

⁷⁹ Companies Act 2006, ss 261-262.

⁸⁰ Andrew Keay and Joan Loughrey, 'Derivative Proceedings in a Brave New World for Company Management and Shareholders' [2010] 2 *Journal of Business Law* 152 - 161, where they discussed potential hurdles a shareholder may face in bringing derivative action such as, establishing prima facie case; implication of company filing response and effect of s263; and possibility of applicability of equitable doctrine of clean hands. See also *Wishart* [2009] CSIH 65; *Franbar Holdings Ltd. v Patel* [2008] EWHC 1534 Ch, on what appears to be conflicting threshold set by the court.

⁸¹ L Linklater, 'Promoting Success: the Company Act 2006' [2007] 28(5) *Company Lawyer*, 129; See also Yap, 'Considering the Enlightened Shareholders Value Principle', 3 (arguing that the inability of stakeholders to enforce directly makes s.172 toothless against directors).

⁸² Peter Loose *et al*, *The Company Director: Powers, Duties and Liabilities* (10thedn Jordan Publishing, Bristol, 2008) 285.

⁸³ Adefolake Adeyeye, 'The Limitations of Corporate Governance in CSR Agenda' [2010] 31(4) *Company Lawyer*, 4-5.

⁸⁴ Bill Perry and Lynne Gregory, 'The European Panorama: Directors' Economic and Social Responsibilities' [2009] 20(2) *International Company*

Perhaps, as submitted, members will only bring action where directors fail to act in good faith to promote the success of the company for the benefit of members as a whole, and, where they fail to act fairly between members,⁸⁵ rather than bringing action for the protection of stakeholders' interests *stricto sensu*. If the foregoing becomes the eventuality, then it means that "parliament has created a right without a remedy which the law abhors".⁸⁶ As noted by a commentator, it appears law makers have "mistakenly encapsulated shareholders' principle into the objective of the companies" in section 172 (1).⁸⁷ This conclusion, it is submitted, is inevitable since stakeholders' interests remain subordinate to shareholders' interests.

Members of the company will be able to assess directors' compliance with the provision of section 172 (1) when business review is tendered.⁸⁸ Although business review has now become a part of financial reporting, the contours of what constitutes an ideal business review remain blurred, and could be amenable to manipulation by the directors.⁸⁹ Indeed, it has been hypothesised that directors may adopt a cynical approach to stakeholders' interests by adopting mechanical compliance with business review regime.⁹⁰ One

and Commercial Law Review, 25. By contrast, shareholders might as well act in 'unenlightened manner' thereby detracting from the substance of the CSR paradigm. See Carrie Bradshaw, 'The Environmental Business Case and Unenlightened Shareholder Value' [2013] 33 *Legal Studies*, 141.

⁸⁵ S.172(1)(f).

⁸⁶ Deryn Fisher, 'The Enlightened Shareholders - Leaving Stakeholders in the Dark: Will S.172(1) of the Companies Act 2006 make Directors Consider the Impact of their Decision on the Third Parties' [2009] 20(1) *International Company and Commercial Law Review*, 10.

⁸⁷ Daniel Attenborough, 'How Directors Should Act when Owing Duties to Companies' Shareholders: Why we Need to Stop Applying *Greenhalgh*' [2009] 20(10) *International Company and Commercial Law Review*, 339

⁸⁸ See s.417(2) CA 2006.

⁸⁹ Michael Page, 'Business Models as a Basis for Regulation of Financial Reporting' (2012). <<http://link.springer.com/article/10.1007/s10997-012-9239-0>> accessed: 25 November 2013.

⁹⁰ Charles Wynn-Evans, 'The Companies Act 2006 and the Interest of Employees' [2007] 36(2) *Industrial Law Journal*, 192.

just hopes that the preceding eventuality does not become a reality.

VIII. Have Regard (amongst other matters) to

The Act provides no definition of ‘have regard to’. It has, however, been suggested⁹¹ that directors should ‘have regard to’ array of other codified duties such as company’s constitution;⁹² creditors’ interest;⁹³ exercise of reasonable care, skill and diligence;⁹⁴ avoidance of conflict of interest;⁹⁵ not accept benefit from third parties;⁹⁶ and disclosure of interest in any transaction.⁹⁷

Keay,⁹⁸ on the other hand, submitted that ‘have regard to’ entails having regard to the interest of constituencies other than those referred to in section 172(1) so far as this promotes benefits to members. Keay’s position has been echoed as being sensible.⁹⁹

The inclusion of ‘among other matters’ means that the catalogues of matters directors should take account of in promoting the success of the company for the benefit of the members as a whole are not exhaustive. This was buttressed by government statements emanating from Lord Goldsmith: “we have included the words, ‘among other things’. We want to be clear that the list of factors [for a director to have regard to] is not exhaustive”.¹⁰⁰ Such ministerial statement, no doubt,

⁹¹ JP Sykes, ‘The Continuing Paradox: a Critique of Minority Shareholders and Derivative Claims under the Companies Act 2006’ [2010] 29(2) *Civil Justice Quarterly*, 5.

⁹² S.171 CA 2006.

⁹³ S.172(3) CA 2006.

⁹⁴ S.174 CA 2006.

⁹⁵ S.175 CA 2006.

⁹⁶ S.176 CA 2006.

⁹⁷ S.177 CA 2006.

⁹⁸ See Keay, ‘S.172 (1) of the Companies Act 2006: an Interpretation and Assessment’, 8.

⁹⁹ Sophia Wesley-Key, ‘Companies Act 2006: Are Cracks Showing in the Glass Ceiling’ [2007] 18(12) *International Company and Commercial Law Review*, 428.

¹⁰⁰ DTI, Companies Act 2006: Duties of Directors: Ministerial Statement (DTI, London, June 2007) 8 at column 846, quoting statement of Lord Goldsmith at Lords Grand Committee, May 9, 2006.

helps in understanding the section, but as noted, it is doubtful if such ministerial clarification will douse the anxiety of directors seeking to avoid falling foul of section 172.¹⁰¹

IX. Employees

Section 172(1) (b) provides for the protection of the interest of the employees as part of stakeholders. There was a similar provision under the Company Act 1980 and Company Act 1985.¹⁰² However, the difficulty of demonstrating breach, and the inability of the employees to directly enforce the right since only the company could bring action rendered the section almost useless.

Similarly, under section 172 (1) the interest of employees is clearly subordinate to the promotion of the success of the company for the benefit of the members as whole. Equally, employees do not have direct enforcement powers except where they double as shareholders. And even then they must seek court approval which stands the danger of being refused. On account of the above shortcomings, it has been argued that section 172 “will not necessarily improve directors’ substantive engagement with employees’ interests”.¹⁰³ In other words, employees’ position is even worse off now than hitherto the case since their interest would have to compete in the same pedestal with other stakeholders’ interests. However, it has been argued that other myriad employees’ statutes can offer protection to employees.¹⁰⁴ It is submitted that the protection afforded to employees by section 172 (1) is inadequate even though they might be protected elsewhere by other legislations. One just expects that the section would not turn out to be nothing but

¹⁰¹ Chizu Nakajima, ‘Whither “Enlightened Shareholder Value?”’ [2007] 28(12) *Company Lawyer*, 354.

¹⁰² See s.309 Company Act 1985.

¹⁰³ Wynn-Evans, ‘The Companies Act 2006 and the Interest of Employees’, 192.

¹⁰⁴ See Kiarie, ‘At Crossroads’, [2006] in Fisher, ‘The Enlightened Shareholders - Leaving Stakeholders in the Dark: Will S.172(1) of the Companies Act 2006 make Directors Consider the Impact of their Decision on the Third Parties’, 11-12.

a mere legislative piece of decoration, with little or no enforcement value, in the same manner as the corresponding section 309 of 1985 Company Act.

X. Creditors

There is a remarkable absence of creditors in section 172(1) (a-f). It can be argued that ‘others’ mentioned in subsection (1) (c) in company of suppliers and customers incorporate creditors, based on *ejusdem generis* canon of interpretation, since they belong to the same genus.

However, a better argument is that creditors were tactically omitted in subsection (1)(a-f) because their interest is already protected by subsection 3; which states that section 172 is subject to any enactment or rule of law requiring directors, in certain circumstances, to consider or act in the interest of creditors of the company. Thus, the circumstances in which directors are supposed to consider the interest of creditors vis a vis shareholders remain uncertain under the Act. Apparently, the efficacy of subsection 3 comes into operation during insolvency or threatened insolvency when directors must cease trading lest they be guilty of wrongful trading¹⁰⁵ and/or misfeasance.¹⁰⁶ Thus, section 172 (1) is subject to Insolvency Act of 1986. In other words, during threatened insolvency directors must consider creditors’ interest which takes priority¹⁰⁷. Indeed, it has been suggested¹⁰⁸ that section 172(3) seems to have preserved such cases¹⁰⁹ that require directors to consider creditors’ interest on imminent insolvency, though owing no such duty to

¹⁰⁵ See s.214 Insolvency Act 1986.

¹⁰⁶ See s.212 Insolvency Act 1986.

¹⁰⁷ Derek French et al, *Company Law* (26th edn Oxford University Press, Oxford, 2009-2010) 485.

¹⁰⁸ Andrew Hicks and SH Goo, *Cases and Materials on Company Law* (6th edn. Oxford University Press, Oxford, 2008) 385.

¹⁰⁹ See *Liquidator of West Mercia Safetywear Ltd v Dodd* [1988] BCLC 250.

creditors.¹¹⁰ It is doubtful whether section 172 added anything new in the interest of creditors.

XI. Teleological Construction of Section 172

Although significant scepticisms underscore the potential of section 172 to cater to stakeholders' interests, nevertheless, there could still be ways of ameliorating the situation. Two options readily come into perspective. Firstly, should lawmakers embark on another legislative reform? Secondly, should reliance be placed on the courts to adopt a pragmatic and purposeful interpretation of section 172 known as teleological construction?

Although initiating another legislative option that would significantly take into account the stakeholders' interests might be a useful policy option, such a wholesome enterprise might not be the optimal option under the current circumstance. First and foremost, it would be unwise to embark on another holistic legislative exercise simply to rectify a perceived anomaly (or misalignment) inherent in one section, which has not even become a subject of substantial judicial interpretation. Secondly, it appears that the lawmakers while craving for greater appreciation of the stakeholders' interests by the directors were reluctant to state so categorically.

In other words, it appears the legislative philosophy behind the enlightened shareholders value might have been to situate the UK corporate governance model somewhere between the traditional common law shareholders' primacy theory and dual consideration theory of some continental European countries; even though some scholars seemingly think otherwise. Thus, the Commission on the Public Policy and British Business suggested that the CLRSG adopts a pluralist approach to corporate governance.¹¹¹ However, this was rejected by the CLRSG on the ground that it might result

¹¹⁰ See *Yukong Line Ltd v Rendsburg Investment Corporation* (No2) [1998] 1 WLR 294.

¹¹¹ See The Commission on the Public Policy and British Business, *Promoting Prosperity: a Business Agenda for Britain* (Vintage, London, 1997) 107.

in directors not being “effectively accountable to anyone since there would be no clear yardstick for judging their performance”.¹¹²

Nevertheless, the CLRSG did not set out simply to codify the traditional shareholders’ primacy theory characteristic of the common law paradigm. Indeed, it was thought that the shareholders’ primacy model should be modernised to incorporate stakeholders’ interests in line with contemporary practice.¹¹³ However, the end result of this preceding legislative effort appears to have been blurred, and even obfuscated by the divergences characterising its understanding. However, these divergences can be settled by the courts through the adoption of a teleological approach to construction of the section.

Teleological construction in this vein entails that the courts construe the section pragmatically to integrate fully the stakeholders’ interest in similar vein as that of the shareholders. One useful means of achieving such a result is for the courts to adopt the approach that actions of the directors that are inimical to the interest of the stakeholders do not promote the interest of the shareholders. The implication is that the directors acting in good faith “to promote the success of the company for the benefit of its members as a whole”¹¹⁴ must take into account the interests of the stakeholders otherwise the resultant actions do not “to promote the success of the company for the benefit of its members as a whole”.¹¹⁵ Arguably, this approach could provide the required interpretative elixir that would lessen the controversy characterising the section. Similarly, it would directly coalesce stakeholders’ interests with that of

¹¹² See Committee on Corporate Governance, Final Report (London: Gee, 1998) para. 1.17, 12 <<http://www.ecgi.org/codes/documents/hampel.pdf>> accessed: 11 February 2014.

¹¹³ See Stephen Girvinet *al*, *Charlesworth’s Company Law* (18th edn, Sweet and Maxwell, London, 2010) 331; OECD, *OECD Guidelines for Multinational Enterprises* (OECD, Paris, 2011) 17-26.

¹¹⁴ Section 172 (1) CA 2006.

¹¹⁵ Section 172.

shareholders. Apart from the fact that the preceding teleological interpretative approach obviates the necessity of further legislative reform; such an approach would further align the UK corporate governance model with that of the significant number of the continental European countries, in line with contemporary practice. Indeed, the divergences that underpin this sphere of corporate governance would continue to hold sway for a while attenuated only by judicial pronouncement and clarification on the real purport of the section.

XII. Conclusion

The enlightened shareholders approach marks a watershed in the protection of stakeholders' interests. Similarly, the scope of constituencies of stakeholders' the directors must have regard to has laudably been widened. Nonetheless, section 172 still suffers from the inability of the stakeholders and even shareholders to directly enforce it, raising doubts as to the potential usefulness of the reform. A stakeholder that doubles as a shareholder and/or shareholder(s) *per se* having the interest of stakeholders at heart might, however, enforce the section which may constitute ameliorating factor to the above shortcomings, subject of course to vagaries of court approval. Indeed, the debate on whether the primacy theory or pluralist theory or even the amalgam of the crystallised theories of both holds sway would, no doubt, continue to dominate the mind of commentators for the foreseeable future. However, courts could lessen the foregoing divergences by adopting a teleological interpretative paradigm that adequately caters for the stakeholders' interests. This preceding could be achieved by a judicial approach premised on the fact that the directors' actions that are inimical to the stakeholders' interests do not 'promote the success of the company for the benefit of its members as a whole'. Apart from the fact that the foregoing approach would streamline the expectations of the stakeholders, such a position would

equally align the UK corporate governance model with the contemporary practice.¹¹⁶

¹¹⁶ OECD, OECD Guidelines for Multinational Enterprises at 17-26.

Case comments

Football must play by the same rules as everyone else – A Comment on *Berg v Blackburn Rovers Football Club & Athletic Plc*

Philip Ebsworth

Background

The case between Blackburn Rovers FC and their former manager, Henning Berg, raised the question of whether the Managing Director of a football club has the authority to enter into negotiations and, subsequently, a contract of employment with a manager on behalf of the football club.

Berg was employed as the manager of Blackburn Rovers FC following the dismissal of Steve Kean in November 2012. The negotiations prior to Berg's appointment were with Derek Shaw, the club's Managing Director, with the owners of the club having no direct involvement. Berg signed a service agreement on 16th November 2012, which Shaw signed on behalf of the club, which would continue until 30th June 2015 and contained a clause which stated that,

“In the event that the Club shall at any time wish to terminate this Agreement with immediate effect it shall be entitled to do so upon written notice to the Manager and provided that it shall pay to the Manager a compensation payment by way of liquidated damages in a sum equal to the Manager's gross basic salary for the unexpired balance of the Fixed Period assuming an annual salary of £900,000.”¹

Blackburn Rovers FC terminated the service agreement with Berg on 27th December 2012 which entitled Berg to the remainder of his salary under the aforementioned clause. However, the club declined to pay the £2.25 million due so

¹ Berg v Blackburn Rovers Football Club & Athletic Plc [2013] EWHC 1070 (Ch), 16.

Berg issued proceedings in February 2013 for breach of contract.

Following the filing of proceedings, Blackburn Rovers filed a formal admission in which they admitted liability to the claim but due to financial difficulties were only able to offer to pay in monthly instalments over a period of four months. However, at the first hearing on 16th April 2013 the club applied to withdraw their admission of liability and instead attempt to defend the claim of breach of contract. Therefore, in the hearing on 26th April 2013, His Honour Judge Pelling QC was required to consider whether Blackburn Rovers had a “*realistically arguable defence*” in order to allow the withdrawal of their admission.

Proceedings

The first defence Blackburn Rovers raised was that the payment due to Berg amounted to a penalty for breach of contract and is consequently unenforceable under English law. HHJ Pelling did not accept this argument and followed the decisions in *Campbell Discount Company Ltd v Bridge*² and *Export Credits Guarantee Department v Universal Oil Products Company*³ which held that payments which are triggered due to a particular event rather than a breach of contract are not to be read as penalty clauses. HHJ Pelling went on to state that once it is established that a party is entitled to terminate the agreement before its expiry, as was the case in this instance, the law relating to penalty clauses is immaterial.⁴

Blackburn Rovers’ main defence rested on the argument that Shaw did not have sufficient authority to bind the club to the service agreement. However, it was very clear from the evidence presented that Shaw did have authority to negotiate and conclude an employment contract with Berg. In his

² *Campbell Discount Company Ltd v Bridge* [1961] 1 QB 445

³ *Export Credits Guarantee Department v Universal Oil Products Company* [1983] 1 WLR 391.

⁴ *Berg* [n1], 34.

witness statement, Shaw stated that the owners had requested him to negotiate a contract with Berg⁵ and that he was the usual signatory on behalf of the club for similar matters.⁶

Blackburn Rovers provided no evidence which suggested that Shaw did not have sufficient authority nor that Berg or his representatives were aware that Shaw's authority was restricted. They attempted to argue that Berg should have been aware that the agreement would have to be authorised by the owners of the club. However, evidence from Olaf Dixon, the Deputy Chief Executive of the League Managers Association, stating that it was normal for a manager to accept that a managing director has the authority to offer the terms they are offering on behalf of the club, strongly suggested otherwise.⁷ Consequently, it appeared that Shaw was not operating outside of the control of the owners but acting within the usual boundaries of normal practice within the football industry.

HHJ Pelling found that Shaw, as Managing Director, had usual authority to make decisions on behalf of the club in the ordinary course of its business as established in *Hely-Hutchinson v Brayhead*.⁸ Therefore it was held that the service agreement was binding on the club and Blackburn Rovers' application to withdraw their admission was dismissed and they were ordered to pay the outstanding sum owed to Berg.

Comment

Following the judgement, the League Managers Association issued a statement stating that:

“It is unacceptable for a Football League club to allow the state of affairs to arise whereby its own case before the High Court is that its Managing

⁵ *Berg* [n1], 36.

⁶ *Berg* [n1], 40.

⁷ *Berg* [n1], 42.

⁸ *Hely-Hutchinson v Brayhead*[1968] 1 QB 549.

Director is out of control and operating outside the authority of the Club's owners."⁹

The strongly worded statement is not unreasonable considering that Blackburn Rovers never really got a legal argument together to support their position. Although it may seem obvious to apply well-established principles of contract and agency law to the case it is an important decision as it further demonstrates that the exceptionally big business of football does not have any exceptions in the law.

The decision in Berg has since been followed in *LNOC Ltd v Watford Football Club*¹⁰ where it was held that a Managing Director has implied authority to do all things which would normally fall within the remit of that office. It further held that there is no need for a formal appointment to the role as long as the board permits or authorises someone to act as Managing Director. This decision gives further weight to the precedent set in *Hely-Hutchinson* regarding actual and implied authority to be equally applicable in the football industry.

It appears sensible to apply agency law principles which have stood for fifty years to the football industry. As football has become a bigger business it has attempted to push the boundaries of 'everyday' legal concepts in order to carve itself out its own law and set of precedents. However the 'everyday' law is the law because it is deemed to be the position and principles that deliver justice to parties involved in a dispute. There is no reason why justice should be administered differently in a footballing context and this decision supports the principle that there is no such thing as sports law, just law applied in a sporting context.

Author's Note: Although the initial reaction is to feel bad for Blackburn Rovers' barrister who was instructed to defend

⁹ League Managers Association, '*Henning Berg Succeeds in High Court Claim Against Blackburn Rovers*' 29th April 2013. Available at <www.leaguemanagers.com/news/news-7195.html> accessed June 2014.

¹⁰ *LNOC LTD v Watford Association Football Club Ltd* [2013] EWHC 3615 (Comm).

what looked like, and was, an unwinnable argument, spare a thought for Derek Shaw who, having instructed his lawyers to admit liability, then had to make a witness statement stating he was a rogue director acting outside his powers whilst at the same time being threatened with disciplinary proceedings for having done so.

Power v Greater Manchester Police Authority¹

Amin Al-Astewani

Background

This case is now part of a number of religious discrimination cases in the last decade which mark the ever-increasing importance of religion and law as a new niche area in the legal arena, particularly in the context of employment. This series of cases can be stretched back to the landmark *Begum*² case in 2006, which then gave way to *Ladele*³, *Chondol*⁴ and *Mcfarlane*⁵ respectively. The most recent addition in this series was *Eweida and others*⁶, which was taken all the way to Strasbourg by the claimants after their claims were dismissed by the UK courts. Since these cases were amongst the first to seriously test the application of newly introduced anti-discrimination and equality legislation, English judges were faced with the difficult task of establishing a precedent for future judges in how such legislation is best interpreted and approached. Religious discrimination cases are often contentious in their very nature, but have also attracted attention because they play into larger debates about the place of religion in the secular courts and the secular public sphere.

The Facts

This case was heard at first instance in the Manchester Employment Tribunal with both the Claimant and Respondent both based in Manchester. Mr Alan Power, an employee of the Greater Manchester Police Authority, brought a claim of unlawful direct discrimination against his

¹ [2010] UKEAT/0087/10/JOJ.

² R (on the application of Begum) v Headteacher and Governors of Denbigh High School [2006] UKHL 15.

³ Ladele v London Borough of Islington [2009] EWCA (Civ) 1357.

⁴ Chondol v Liverpool City Council [2009] UKEAT 0298/08/JOJ.

⁵ *Mcfarlane v Relate Avon Ltd* [2009] UKEAT 0106_09_3011.

⁶ *Eweida and others v UK* [2013] ECHR 37.

employer contrary to the provisions of the Equality (Religion or Belief) Regulations 2003. These regulations, now incorporated into section 19 of the Equality Act 2010, state that direct discrimination is established on the grounds of the religion or belief of an individual if “A treats B less favourably than he treats or would treat other persons”.

Mr Power was a committed member of the Spiritualist Church and a keen believer in spiritualism, life after death and the use of mediums and psychics to contact the dead. He successfully applied for the post of a Special Constabulary Trainer in August 2008. Shortly after, his employer received statements from two police officers regarding his inappropriate conduct as a volunteer in neighbouring police forces. An investigation was consequently mounted, which found that Mr Power had also been providing other police officers with a CD-ROMs and posters related to his personal beliefs. As a result Mr Power was dismissed after only three weeks in his employment. A crucial part of the dismissal statement referred to his “work in the psychic field” as one of the major reasons for his dismissal, which is particularly what prompted Mr Power to complain of discrimination on grounds of his religion or belief.

The decision

The tribunal at first instance noted that the Spiritualist Church, of which Mr Power was a member, was established in 1853 and that spiritualism was listed as the eighth largest faith group in Britain with 32,404 adherents in the 2001 census. It considered that this fulfilled the test for a religious or philosophical belief capable of being protected under the 2003 regulations, as set out in *McClintock*⁷: it had sufficient cogency, seriousness, cohesion and importance, and was worthy of respect in a democratic society. However the tribunal decided that Mr Power had not been discriminated against on the grounds of those protected beliefs, but found he had been dismissed for two separate reasons. The first

⁷ *McClintock v Department of Constitutional Affairs* [2008] IRLR 29.

was his inappropriate conduct as a volunteer which highlighted he was not a suitable candidate to train young police officers. The second was his distribution of posters and CD-ROMs around the police station advertising his personal beliefs about spiritualism which did not meet the standard of professionalism required in the work place. Mr Power appealed the decision.

The EAT headed by Judge Peter Clark affirmed the Tribunal's use of the classic *Igen*⁸ approach where a reverse burden of proof is imposed. In *Igen* the Court of Appeal applied a two-stage approach, in which the claimant first has to establish that there has been unlawful discrimination. The burden of proof then shifts to the employer to prove otherwise; if their explanation is not adequate, the tribunal has to find the discrimination proven. The EAT endorsed the Tribunal's decision that Mr Power did validly raise a *prima facie* case of unlawful discrimination. It accepted however that the explanation given by the Respondent was also valid. Direct discrimination can of course not be justified unlike indirect discrimination, thus the only reason the respondent's explanation was valid was because they proved that the reasons for dismissal were not related to the beliefs of Mr Power. This was arrived at using the "reason why" question postulated by Lord Nichols in *Shamoon*⁹ and adopted by Elias P in *Ladele*. It was also used in *Chondol* and *Mcfarlane*. This question focuses on the reason for a dismissal and enables a tribunal to make a distinction between treatment complained of on the ground of the Claimant's religion (unlawful) and on the ground that the Claimant was improperly foisting it on service users (lawful). Thus the EAT concluded that "the reason for the Respondent's dismissal of the Claimant was, in part, the manifestation of his spiritualist beliefs in the material which he had earlier distributed, not the fact of his beliefs.

⁸ *Igen v Wong* [2005] ICL 931.

⁹ *Shamoon v Chief Constable of the Royal Ulster Constabulary* [2003] ICL 337, 341-2.

Accordingly, there is no error of law disclosed in this appeal and it must be dismissed”.

Comment

It is welcoming to have a Manchester case which sets two significant precedents in this area of the law. The first precedent is in the definition of ‘Religion or Belief’. This case has further confirmed the wide scope of the definition by holding that a belief in spiritualism and the philosophical belief in life after death and psychic powers falls within the Employment Equality (Religion or Belief) Regulations 2003. The EAT in *Grainger*¹⁰ had previously concluded that an asserted belief in man-made climate change, together with the alleged resulting moral imperatives arising from it, was also capable of constituting a ‘philosophical belief’ for the purpose of the 2003 Regulations because it met the criteria laid out by the Article 9 jurisprudence of the European Convention on Human Rights (ECHR) which was directly relevant. This is welcoming news for employees who may face discrimination because of certain beliefs they hold, whether religious or philosophical. The second significant precedent is that this case affirmed the distinction that judges will draw between treatment on the grounds of a person’s belief and on the grounds of the manifestation of those beliefs. This will be a welcoming development for employers as it sets quite a large hurdle for claimants to pass. Although I believe the decision was correct in this case because of the clear reason of misconduct by the claimant, I am uncomfortable with this second precedent in the way it sets such a large hurdle for employees. The distinction is not very easy to draw, and it is questionable whether discriminating on grounds of religion require you to be specifically motivated by that person’s religion. This is indeed why most direct discrimination claims have thus far been unsuccessful, failing because the actions of the defendant are found not to be on the grounds of religion.

¹⁰ *Grainger PLC v Nicholson* [2009] UKEAT 0219/09/ZT.

Harry Street Lecture
20 November 2013

The Queen and the Constitution (and Me)

Rodney Brazier

Harry Street was on the committee that appointed me here in 1968. When it was his turn to interrogate the 21 year-old, third-year, undergraduate, it was suggested that he ask me about my constitutional law dissertation. Harry simply replied “I don’t know anything about that. I’ll ask questions on administrative law.” And they were tough. But I bluffed my way through somehow. Harry was to give an enormous early boost to my career. We co-edited the late Stanley de Smith’s classic work, *Constitutional and Administrative Law*. Ever keen to enhance his salary, I can still see Harry’s beaming face as he told me that, in its first year, and throughout the world, it had sold a staggering 10,000 copies.

Harry was the leading public lawyer of his generation. In every way it was a privilege to have known him. That view is shared by thousands of his students and his colleagues. You will understand, therefore, how moving it is for me to give this lecture in his memory.

One February morning in 1952, a 5-year-old boy was daydreaming in his state primary class. Suddenly, the head teacher came in. He said that the King, George VI, had died, aged only 56. Princess Elizabeth was our new Queen. She was only 25. I was that little boy. To me, this news was of no interest at all. I couldn’t wait for break-time, so that we could resume our snowball fights. That working-class scallywag was to become a frequent visitor to Buckingham Palace, discussing constitutional matters. More of all that later. (In case I forget to say, though, my first-ever visit there was at age 8. My parents had taken me up to London to see the sights. Having watched the changing of the guard at the Palace, I desperately needed to do what little boys often

desperately need to do. My Mum whizzed me around the corner to Constitution Hill, where I did the necessary discreetly by the wall.⁵⁹ years later, I publicly apologise. I know that the Queen would understand.)

But I digress. Like many of her predecessors, the Queen handwrites a daily diary. What an incomparable mine of information that must be! But, following precedent, none of it will be published until after her death. Even so, enough is known after 61 years to assess the effect of the Queen on the British constitution, and vice versa. My analysis today will have to be very selective. The full story merits a new book. It will be interesting - I hope! - to see exactly what has happened since 1952, why it has occurred, and whether it was planned.

Monarchy is a unique part of a constitution. Unlike all the other parts it consists of one person. So it's impossible to think about the monarchy without picturing the Queen. She fulfils her duties in much the same way as she always has. Even republicans acknowledge that, if there must be a monarchy - which they reject - the Queen has performed well. We also picture the next King - the Prince of Wales - and the King after him - the Duke of Cambridge. (Perhaps we can leave Prince George out of all this for now, gurgling happily away.)

What did the Sovereign's constitutional powers look like on the death of George VI, on 6 February 1952? Queen Victoria (who reigned from 1837 to 1901) had used her constitutional powers fully towards her Prime Ministers. She helped her favourites. She obstructed Ministers for whom she did not care, or of whose policies she did not approve. But even Victoria had to accept, by the latter part of her reign, that the monarchy had changed. By that time, the Government represented the wishes of the (albeit small) electorate. The vote had been given to more and more categories of men. That (albeit limited) democratic legitimacy, possessed by the Government of the day, allowed two constitutional conventions to develop. One is what I call for shortness "the cardinal convention". It prescribes that royal

action will normally only take place on ministerial advice, even if a Sovereign personally dislikes that advice. That convention is counterbalanced by what I dub “the tripartite convention”. It was identified by Walter Bagehot as the constitutional right of the Sovereign to be consulted, to advise, and to warn Ministers. (Those labels of mine were used in the continuing case of *Evans v. Information Commissioner*.) Thus even Victoria came to accept that she must normally act as elected Ministers advised her, even if it was uncongenial to do so. In return she could offer views – often trenchantly – in private to those Ministers, confidentially trying to steer, perhaps to help, Governments of any political party. That position was solidified by her successors, Edward VII, George V, and George VI. (We can ignore the 11-month reign in 1936 of Edward VIII. He is remembered today only for his Abdication. How lucky Britain was to be shot of both him and Mrs Simpson.)

All a Sovereign’s public speeches (except for the Christmas Day broadcast) are approved in advance by Ministers. Apart from those speeches, the Queen has been notable for her public silence, just as she had been as Princess Elizabeth. She has never given an interview, or published anything. We have no idea what the Queen *personally* thinks about public issues, politics, the political parties, or indeed very much else. (That has not prevented self-styled “royal experts” popping up on TV, with monotonous regularity, talking with confident authority about things of which they can know nothing.) The Queen has maintained that self-denying ordinance since 1952. (By contrast, much has been heard from the Prince of Wales. That raises intriguing constitutional questions, and there might be difficulties ahead for him. But time prevents me from addressing them here.) Perhaps the Queen’s silence resulted both from her youth in 1952, and the reticence of the times. In those days it just wasn’t done for members of the royal family to inflict their opinions on the public. Every attempt by the media to persuade the Queen to depart from

her public silence has failed. But so what? Why have I gone off on this apparent tangent?

This royal silence has underpinned both the cardinal convention, and the tripartite convention. The Queen has adhered scrupulously to both. Her inscrutability has enhanced them. A Sovereign must be apolitical, and be seen to be apolitical. To have no publicly-known views about any significant matter enhances the perception that she is above and beyond party politics. It is not known how she might vote - if she had a vote. Obviously, the Queen *must* have her own, personal, views about current affairs and public policy. How could she not? All her Prime Ministers have attested to the value which they have obtained from her views, expressed in complete confidence at their weekly meetings ("audiences"). They take place between the Sovereign and the Prime Minister alone. No minutes are taken. That enables complete frankness on each side. All her Prime Ministers have acknowledged the Queen's great knowledge of public affairs, of policy details, of other heads of state and government, and not least of the Commonwealth. It is at these audiences that the tripartite convention mainly operates. The Queen's public silence has vastly strengthened it.

By 1952, and despite the cardinal convention, real royal power remained. Thus it was still for the Sovereign personally to choose a Prime Minister, except, of course, after General Elections where one party emerged triumphant. Moreover, it was for the Sovereign to decide whether to grant a dissolution of Parliament. Such decisions could be difficult in certain political circumstances. And the royal prerogative itself - that vast collection of common-law powers - remained largely in place, ill-defined, never having been officially catalogued. Thus at her Accession in 1952 the Queen inherited a monarchy with real constitutional powers. And please don't assume that these were unused, or unimportant, powers. Her granddad, George V, had had to choose the Prime Minister in 1923, 1924, and 1929. And he had taken a major role in 1931 in the formation of the coalition, so-called

National, Government. The Queen's dad, as late as 1951, was vexed by the prerogative of dissolution. The 1945 Labour Government had squeaked back to office at the February 1950 Election, with a tiny majority. There was much public debate about whether the Prime Minister, Clement Attlee, would be entitled to a further dissolution whenever he asked for it. (In fact, he obtained an Election in October 1951. Labour lost to the Conservative Party under Winston Churchill.) And just three years into her reign, the Queen had personally to choose a new Prime Minister. As it turned out, she had to choose the two after him as well. All three happened to be Conservatives. I should like to look briefly at those choices.

Unlike the other political parties, the Conservatives had no formal leadership election system until 1965. (I'll say more on that later). The choice of a successor to the great Churchill was easy. He retired, aged a mere 80, from his peacetime premiership, 1951 to 1955. Sir Anthony Eden had long been his unrivalled heir apparent. In less than two years, however, Eden had to resign through illness in the wake of the Suez fiasco. The succession in 1957 lay between two Cabinet Ministers, Harold Macmillan, and R A Butler. Very limited consultations, conducted by the Queen's Private Secretary, resulted in the Queen summoning Macmillan to the helm. There were rumblings from the losing side. (By the way, I once saw Macmillan, from the public gallery, in action at Prime Minister's Questions in the House of Commons. He was a class act.)

Six years later, in October 1963, Macmillan was struck by what he later called "the stroke of fate". He (wrongly) thought that he had prostate cancer, and decided to resign. This time there were *several* rival successors. The scene was set for what some commentators have described as the Queen's worst constitutional mistake. Briefly, this is what happened. From his hospital bed Macmillan initiated what he was pleased to call the "customary processes of consultation" throughout the Conservative Party. (In truth, there were no such "customary processes".) There was, however, an unprecedentedly-wide

canvass of Conservative opinion. The Cabinet, junior Ministers, Conservative MPs, peers, and others, were asked their views, by various specified individuals. Macmillan collated the results into a memorandum, which he read to the Queen when she visited him in hospital. The recommendation was that the Foreign Secretary, the 14th Earl of Home, should become Prime Minister, which he did. (I heard the news at school.) This was very surprising - and highly controversial. The "customary processes" were damned by some for containing leading questions, designed to do-down a leading candidate, R A Butler. (Poor old Butler again, losing for a second time. He never was PM.) The Conservatives - not surprisingly - decided that such an exercise must never be repeated. And so in 1965 they fell into line with the other parties by adopting a clearly-democratic, secret, balloting system for their leader. (By the way, I saw Home as Prime Minister during the 1964 General Election. He was addressing an open-air crowd, in a public car park, open to all. Simpler, safer days! In the same campaign, I listened to the young Labour leader, Harold Wilson, making a public speech, in our small public hall in Loughton, in Essex. (No "Towie" jokes, please.) He was accompanied by an elderly, and silent, Clement Attlee.)

The criticism of the Queen is that she was constitutionally wrong to rely so heavily on Macmillan in 1963, and that she should have consulted more widely. On one point, all are agreed. The advice that Macmillan gave wasn't *ministerial* advice. His written resignation had been delivered to the Queen earlier that day (by his Principal Secretary - he didn't rely on Royal Mail). So whatever he said to the Queen could not amount to ministerial advice. Apart from that, though, I think that the criticism of the Queen is misplaced. Was the Queen *really* expected to set aside a long document, apparently based on rigorous, widespread consultation within the governing party? How could her Private Secretary possibly have rivalled the Conservative Party's consultations? I can't see how the Queen could have acted differently, without being seen to meddle in a political party's

internal affairs. In any case, after 1965 similar circumstances could not recur.

Happily for democracy, a new convention about prime ministerial succession emerged afterwards. It governs the Macmillan-type situation. The convention is that, in such a case, the resigning Prime Minister is expected, if possible, to remain in office until the Government party has elected a successor. The winner will then be appointed formally by the Queen to office, following the resignation of the outgoing Prime Minister. That process was followed after Harold Wilson's surprise announcement in 1976 that he would resign after his party had elected a new leader. It was repeated in 1990 when Margaret Thatcher was ditched by her party, and again in 2007 when Tony Blair retired after 10 years as Prime Minister. That convention marks a significant reduction in royal responsibility, and in my opinion rightly so. The days of new Prime Ministers being personally chosen by the Sovereign (such as Eden, Macmillan, and Home) have ended in the Queen's reign.

But wait a minute. What about the Queen's role when a hung Parliament is elected? After inconclusive elections in 1923 and 1929, George V had to make real, personal, choices of Prime Minister. There was a hung Parliament in February 1974 - the first of the Queen's reign. In 1974, the politicians sorted out what should happen. The Queen remained constitutionally inactive. A minority Labour Government took office. By 1974 it had become accepted that such a highly-charged political decision should be made by politicians, not by the unelected head of state. (Please note, though, that the royal prerogative remains intact. *If* in such circumstances the politicians failed to sort it out, only the head of state could do so. Only the Sovereign has the legal power, under the royal prerogative, to appoint a Prime Minister. Perhaps, though, this will never have to happen.)

We owe it to the constitutionally-ambitious Gordon Brown (Prime Minister 2007 to 2010) for the codification of that, and other, processes. He asked the then Cabinet Secretary, Sir Gus O'Donnell, to produce what was to be

published as the *Cabinet Manual* 2011. (Incidentally, I have met Lord O'Donnell, as he now is, on a number of occasions. (And incidentally, incidentally, have you noticed my First Law of Name-Dropping? If you're going to drop names into the conversation, make them good ones. I'm still waiting for my chance to say to someone, boring away about this or that, "I do agree, and in fact said just that at dinner yesterday to the Queen". I'm digressing again!) Anyway, on my first visit to meet Gus O'Donnell at the Cabinet Office, I was keen to visit its loo. I didn't need to use it, you understand. I just wanted to check a rumour. I had heard that each and every sheet of loo paper there bears the solemn declaration "Property of Her Majesty's Government". What joy it would have been to use such paper if you happened to dislike Her Majesty's current Ministers! Alas, however, it's all an urban myth.) That document (the *Cabinet Manual*, not the loo paper) reduces existing non-legal rules and practices to writing. Chapter 2 explains how Governments come to office, including from a hung Parliament. That chapter was published, in draft, *before* the May 2010 General Election. (I had a part in drafting it.) Its procedures were - broadly - followed after Mr Brown lost that election. The present Coalition Government is the result. The Queen discreetly remained out of the way at Windsor throughout the Conservative - Lib Dem negotiations. Mr Brown actually resigned before those talks had finished. As a result, Mr Cameron had to tell the Queen at his first prime-ministerial audience that he didn't know whether there would be a Coalition, or a minority Conservative, Government. But what *is* critical is this. In May 2010 the political dance was performed to music scored in the then draft *Cabinet Manual*. The *politicians* resolved the electoral stalemate - and for the sake of democracy, quite right too. A Sovereign's role in government-formation has, therefore, shrunk vastly since 1952.

The monarchy has been changed by that Coalition Government. First, the Fixed-term Parliaments Act 2011 provides that Parliament lasts for a fixed, five-year

period. Ever since a King first summoned the first-ever Parliament, the Sovereign has had the prerogative power to dissolve it at any time. That Act abolished it. We know that the next election will be on Thursday, 7th May 2015. Put that date in your diaries. And then every five years after that. There could only be an “early” Election if MPs were to vote for it, according to rules in the Act. Before that statute, Prime Ministers asked for Elections when their party was doing well in the opinion polls – *quelle surprise!* They had an unfair advantage over the Opposition. Now, though, election dates are not a matter for the Prime Minister or any concern of the Sovereign. May we not assume that the Queen, on behalf of herself and her successors, might well have said privately, as someone did in *Hamlet*, “For this relief, much thanks”?

Secondly, and as everyone knows, the rules about succession to the Throne are being changed. The Succession to the Crown Act 2013 will be brought into force when all the Commonwealth realms have approved its provisions. (They have already agreed the principles in that Act.) It does three things. First (and at long last) sex discrimination in the line of succession will be abolished. At present, males have preference over females. If the Duchess of Cambridge’s baby had been a girl, then under the new Act any subsequent boy would not, as it were, shove his older sister aside and take her place in the line of succession. (But of course their first child was a boy. Prince George is next in line to the Throne after his dad.) Secondly, the rule which bars the Sovereign from marrying a Catholic, and which automatically removes from the succession anyone who marries a Catholic, will be repealed. The Sovereign, however, *must* still be a communicating Anglican. Finally, the idiotic Royal Marriages Act 1772 will be replaced. That Act requires *anyone* in the line of succession – however remotely – to obtain the Sovereign’s consent to marry. Without it, any such marriage is void. Perhaps if you were to check, one of you might be (say) 767th in that line. If you plan to marry before the 2013 Act comes into force, you must ask for a lovely document, signed by the Queen, to say that she consents. If you don’t, your

marriage would be void. Surprisingly (at least to me) Buckingham Palace has many such requests every year from careful people. The 2013 Act will require only the first six people in the line of succession to seek the Sovereign's consent. If any of them did not (which is unlikely) the subsequent marriage itself would be valid, but the miscreant could never become the Sovereign.

The Queen's father would be astonished by the constitutional changes made to the monarchy over the past 61 years. The Sovereign has become much more of a ceremonial and representative head of state. But the tripartite convention, for example, ensures the continuation of a real, though limited, constitutional contribution by the head of state.

That review of more than 60 years raises at least two questions. First, was this reduction in royal power all planned? I am as sure as I can be that the answer is "No". The British constitution evolves, and changes, in fits and starts. The monarchy, as part of that constitution, is no different. We can see a reason for each, specific, change in the Queen's position. Thus, for example, her discretion to appoint a Prime Minister was limited by the Conservatives' adoption of leadership election rules; and so on. There was - is - no master plan, to my knowledge, and in my judgement. Secondly, what is the Queen's view of these changes? Has she privately welcomed them, resisted them, embraced some, but not others? I don't know. But isn't it a reasonable inference that, in the main, an unelected monarch, fully attuned to the nuances of the constitution, would welcome politicians taking responsibility for their own, political, decisions? We shall have to wait for the publication of the Queen's diary to find out her opinion about all this. But, if we combine her robust good health, the longevity of her mother (who lived to be 101), and the time needed after the next Accession to make that diary available, we mustn't hold our breath to read it.

Those of you who are still awake might be thinking that, while I've droned on about the Queen and the constitution, I

haven't said very much about the "me" in the title of this lecture. I have published a lot about the monarchy. No doubt some people say that I write of little else - although even they must agree that it's all frightfully good. But I've also enjoyed more personal experiences. They all began purely by chance. Late in 1995, on a whim, I sent a copy of an article of mine to the Queen's Private Secretary. In his delightful thank-you note Sir Robert Fellowes asked whether, the next time I was in London, I might like to call in to Buckingham Palace for a chat ("call in"!). Well, I carefully pondered that query for all of 30 seconds, and decided that I would go to London as soon as I decently could. I first entered the Palace in January 1996. Why had he asked to see me? I must explain what the Queen's Private Secretary does. He - always a he, so far - is the Sovereign's most senior, permanent, adviser. Very briefly, and of relevance here, he advises the Sovereign on her rights and duties in relation to the constitution, the Government, Parliament, and the Commonwealth. He must keep abreast - from various people and places - of political and public policy issues, and of changes in constitutional law and practice. Of course, he is entirely impartial and politically neutral. Very kindly, Sir Robert wanted to feel able to consult me from time to time about constitutional issues, whether real, possible, or theoretical. All this would, of course, be in strict confidence. And so it all began. Sometimes meetings take place at Buckingham Palace, sometimes elsewhere. There have also been many written exchanges, usually initiated by me, in the form of unsolicited memorandums, and very occasionally phone calls, usually to me.

Have my views, opinions, thoughts, suggestions, memoranda, jests - for there is always fun at Buckingham Palace - actually made any difference? Here I am in a difficulty. What has passed between me and Buckingham Palace (and indeed more widely in government) is, and always will be, completely confidential. The relationship couldn't exist on any other basis. But let me say this. Possibly now and again my opinion on some constitutional matter -

given as impartially as I can - might confirm a view already held at Buckingham Palace, and I suppose that may be useful. I can't recall a Private Secretary ever saying to me, "That's a good idea - we'll do that". Rather, he absorbs my views, and mulls on them later, along with all other considerations. And that's all I'll say about that. Save to add this. I once had to postpone a seminar because it clashed with a visit to the Palace. The very next day after my visit, the official, public, announcement was made of William and Kate's engagement. The rumour spread around the Law School that I had gone to the Palace to confirm that this would be OK from a constitutional point of view. Sad to say, in fact the announcement was as much news to me as to any other member of the public.

To end, I return to the beginning. Since 1952 the monarchy has slowly altered. That has enhanced parliamentary democracy. Formal royal power, still possessed by the Queen's dad, has been reduced. The Sovereign in 2013 retains some constitutional authority. This is seen in particular in the tripartite convention. And the Queen *might* one day have to be the final arbiter if ever politicians could not resolve some constitutional crisis. Now there is, and has been for a long time, a case that the United Kingdom should become a republic. That case is logical and weighty. But it has very little popular support while the Queen is on the Throne. What might happen to the republican cause in the future must wait for the Accession of the next King, many years ahead. I wonder where, and how old, I will be when I hear that news.

