Northern Ireland Legal Quarterly

Volume 70 Number 4

Winter 2019

Chief Editor

Dr Mark Flear

Queen's University Belfast

School of Law

NILQ

Contents

ARTICLES

The political legitimacy of company law and regulation
Daniel Attenborough **383**

Theses on the relationship between rights and social struggle
Dimitrios Kivotidis **407**

Causing controversy: interpreting the requirements of causation in criminal law and tort law
Gemma Turton and Sally Kyd **425**

Charitable purposes, demonstrable benefit and the role of the Charity Commission: the fourth *Pemsel* head recast
Lara McMurtry **445**

The Stephen Livingstone Lecture: 'The problems with human rights'
Brice Dickson **467**

Barriers to High Court appointments in Northern Ireland
John Morison, Brice Dickson and Andrew Godden **479**

NOTES AND COMMENTARIES

The Irish Supreme Court judgment in *Nano Nagle School v Marie Daly*: a saga of litigation
Shivaun Quinlivan and Charles O'Mahony **505**

Liability threshold for damages in public procurement: the EFTA Court's *Fosen-Linjen* saga
Albert Sanchez-Graells **511**

ISSN 2514-4936 (online) 0029-3105 (print)

NILQ 70(4): 383–405

The political legitimacy of company law and regulation

DANIEL ATTENBOROUGH*

Durham Law School

Abstract

Two interrelated objectives are pursued in this article: the first concerns the relationship and interaction between the UK's company law and market-based regulation; the second explores the legitimacy of market-based regulation in light of its potential to mould and influence the substantive law. Although the article finds that the two systems retain carefully defined, essentially consistent and mutually complementary roles, it submits that market-based formations run the risk of not readily or plausibly lending themselves to dominant political and democratic accounts, which are deployed customarily to substantiate the legitimacy of state interventionist techniques. Simultaneously, the deployment of a rival conception of legitimacy, conceived as technocratic expertise and market consensus and conformity, is problematic for a number of theoretical and practical reasons. This has implications for the effects and outcomes of the UK's company law and governance.

Keywords: company law; corporate law; market regulation; regulatory theory

1 Introduction

Two interrelated objectives are pursued in this article: one instrumental, the other substantive. The narrow objective is to address a question that has attracted very little attention in company law scholarship,[1] although it is, it must be said, of considerable theoretical, and potentially also some practical, importance in respect of the direction and control of companies listed on the London Stock Exchange. The question concerns the interaction between substantive company law and market-based regulation. Does UK company law and market-based regulation work together as, in some sense, a single coherent, integral body of rulemaking, or do they constitute two separate entities, two streams running on parallel lines, one of which occasionally feeds the other, but which are

* Associate Professor in Corporate Law, Durham University. Earlier versions of this article were presented at the 2016 Annual Conference of the Society of Legal Scholars at the University of Oxford, and at seminars in Manchester University, University College London and University of California, Berkeley. I am thankful for the extremely helpful comments and suggestions received from participants at these events, and especially to Marc Moore, Paul O'Connell, Gavin Phillipson, Chris Riley and Alexander Williams. The usual disclaimers apply. Also, a significant part of this work was undertaken while I was at the University of California, Berkeley, to which I am grateful for providing a productive and stimulating working environment.

1 Virtually the only partially related discussion of which the author is aware is in D Kershaw, 'Corporate Law and Self-Regulation' in J N Gordon and W G Ringe (eds), *The Oxford Handbook of Corporate Law and Governance* (Oxford University Press 2018) 869.

destined for ever to retain their separate roles? In posing this question, it must be made clear that we are not concerned with the question of how well law and regulation fit together, but with the logically prior question of *how* they fit together. It is trite that large areas of both the law and market-based regulation are problematic, but is it one problematic area or are there two? This may seem a rather theoretical, or even metaphysical, question. Yet, there are a number of important practical issues involved here, in the sense that market-based regulation is a dynamic, developing body of rules and norms just as much as legal doctrines and legal decisions are, and the way in which they map onto each other becomes a matter of no little importance.[2] The most empirically significant regulatory formations on which we focus are: the Corporate Governance Code, the Listing Rules and Listing Requirements, and the Takeover Code.

While there is in one respect no neat dividing line between substantive law and market regulation,[3] the two species of rules remain distinguishable.[4] First and foremost, companies' legislation and the courts today function to generate higher order rules that set out a template for the processes and constraints of the formation and continuing regulation of companies during their lifecycle. Second, over the past several decades substantive company law has, in different shapes, focused on managing and containing the problem of managerial agency costs: the economic costs incurred when managers act in their own, not the shareholders' interests. Notwithstanding, in targeted areas of socially sub-optimal behaviour and outcomes the longstanding British regulatory style has created a presumptive view that market participants are entitled to generate and enforce lower order regulation to resolve the problems of their own making. Even where the state *is* actively responsible for rule design and enforcement, sometimes the law has remained curiously indeterminate. On this basis, market regulatory forms have the potential to mould and influence the substantive content of that law. Properly understood as facilitating productivity organised in companies by reducing suboptimal market behaviour and managerial agency costs, these regulatory devices purport to maintain an exclusive focus 'on the (functional) means rather than (normative) ends of purportedly efficient boardroom practices'.[5] However, the regulator constituencies tend in practice to embody a 'common investor-protectionist ethos, and corresponding disregard for public policy concerns extraneous to considerations of shareholder welfare'.[6] Although this practical

2 M Petrin, 'Regulatory Analysis in Corporate Law' (2016) 79(4) Modern Law Review 537, 529. See also, P L Davies, 'Corporate Boards in the United Kingdom' in P Davies, K Hopt, R Nowark and G von Solinge (eds), *Corporate Boards in Law and Practice: A Comparative Analysis in Europe* (Oxford University Press 2014) 713, 773.

3 An early example that covers the main arguments along jurisprudential lines can be found in J Raz, 'Legal Principles and the Limits of Law' (1972) 81(5) Yale Law Journal 823. For a more recent example focused specifically on corporate regulation, see E Ferran, 'Corporate Law, Codes and Social Norms – Finding the Right Regulatory Combination and Institutional Structure' (2001) 1(2) Journal of Corporate Law Studies 381, 382 and 386–390.

4 It is submitted that the ideological conservatism, enumerated by the likes of Ferran (n 3), can be viewed as symptomatic of a discipline trying to face up to a new and rapidly changing scene. Conversely, there is a genuine intellectual concern, forcefully articulated by Simon Roberts, that lumping together state law and diverse forms of delegated law obscures the distinctiveness of centralised forms of governance and lets blinkered lawyers in to view other forms of normative ordering through distorted lenses. On this point, see S Roberts, 'Against Legal Pluralism: Some Reflections on the Contemporary Enlargement of the Legal Domain' (1998) 42 Journal of Legal Pluralism 95; S Roberts, 'After Government? On Representing Law without the State' (2005) 68 Modern Law Review 1. See also, W Twining, *General Jurisprudence: Understanding Law from a Global Perspective* (Cambridge University Press 2009) 371–375.

5 M Moore, Corporate Governance in the Shadow of the State (Hart 2013) 172.

6 Ibid. This is discussed further in section 3.

understanding includes and is consistent with company law's shareholder *prioritisation*, a question mark is raised about its compatibility with company law's shareholder *weighting*. Specifying, market-based regulation portends to structure and amplify shareholder interests and value pressures (relative to the UK's substantive company law-making bodies), which risks further reducing management insulation from shareholder value pressures.

Typically, the way in which companies are governed, and the way such governance is perceived, contributes inexorably to the political legitimacy of corporate control, and this legitimacy underpins both the acceptance and the effects of corporate activity. From the perspective of orthodox political narratives and democratic criteria, Parliamentary law-making derives a substantial part of its procedural and substantive legitimacy from sovereign prerogative, processes of consensual democracy, and accountability to the courts. Similar thinking can be applied to the court's role, which is to apply the prescriptions of the legislature, or the established principles deduced from a series of precedents, to individual disputes. Yet, an equivalent narrative for today's market-based regulation is less straightforward when it encounters and interacts with state-legitimacy pressures. These dynamic and semi-autonomous rule-systems, which are often thought to derive survival and legitimacy from rival notions of 'neutral-technocratic' expertise and market consensus,[7] tend to reclassify acts conventionally regarded 'public' and 'political' as 'commercial', 'private' and, therefore, properly 'legal'.[8] This approach leads the article to a second, and broader, question: does market-based regulation retain a firm basis on which to rest its continuing legitimacy? The obvious tension in this regard is that 'these authorities, or some of them at least, wield significant powers which may be abused to the detriment of their own members, third parties or the public at large: they may be exploited, or their interests may be otherwise insufficiently taken into account'.[9] This article foregrounds this tension between company law and market-based regulatory legitimacy and examines the strength of the case for the possibility of market-based regulation obscuring the basic principles of law around which that regulation has developed.

2 The role and limits of UK company law

In order to make sense of the question of legal-regulatory 'fit', it is necessary to identify the content and purpose of the UK's company law. Functionalist accounts of a particular set of legal rules focus typically on the purposes served for society by the rules in question.[10] In this regard, company law's role is to provide continuing 'rules of the game' to facilitate and control the activities of business associations during their lifecycle. A functionalist understanding, accordingly, seeks to explain how the system of company law achieves this end.[11] First and foremost, its legal framework primarily enables the structuring of economic power of businesses that 'incorporate' to secure such advantages as convenience, financial flexibility and limited liability.[12] For example, company law

7 Moore (n 5) 170–171.

8 F Vibert, *The Rise of the Unelected: Democracy and the New Separation of Powers* (Cambridge University Press 2007) 9–10.

9 A C Page, 'Self-Regulation: The Constitutional Dimension' (1986) 49(2) Modern Law Review 141, 142.

10 J Armour, 'Who Should Make Corporate Law? EC Legislation Versus Regulatory Competition' (2006) 58(1) Current Legal Problems 369, 371–373. For a similar account of US corporate law, see M A Eisenberg, 'The Architecture of American Corporate Law: Facilitation and Regulation' (2005) 2(1) Berkley Business Law Journal 167, 169.

11 Armour (n 10).

12 See e.g. R H Coase, 'The Nature of the Firm' (1937) 4(16) Economica 386.

establishes the structure of the corporate form, which acts as an incentive to organise productivity and plays a critical role in promoting enterprise and investment. Other examples, of the many in which the law designs basic 'nuts and bolts' of productivity organised in companies, include: the corporate constitution, corporate governance issues (including the directors' role, duties and responsibilities), as well as shareholder remedies, the protection of creditors, and directors' liabilities on insolvency. As a second-order objective – and one that flows from the use of the corporate form – company law plays an important role in mediating the principal–agent conflict between the company's shareholders and its hired directors. This assures that the directors are responsive to the shareholders' interests rather than pursuing their own personal interests.[13] Minimising agency costs is achieved through regulatory strategies that prescribe substantive terms that govern the content of the principal–agent relationship, tending to constrain, either directly or in practice, the directors' behaviour[14] and governance strategies that seek to facilitate the shareholders' empowerment and control over managerial behaviour.[15] Ultimately, the UK's legal ecology directly or indirectly assigns priority interest to the company's shareholders relative to other affected parties in the effective functioning of wealth-generating companies.

Notwithstanding, the UK corporate legal order does not address every aspect of business activity. In particular, the government has, as a historical matter, sometimes considered it more expedient to defer to market mechanisms to constrain opportunistic behaviour, excessive risk-taking and other socially undesirable behaviour.[16] When the substantive law must address, and historically has addressed, the doctrinal minutiae of business regulation and, in particular, the radical or game-changing questions central to company law and policy, its rules and decisions are not, of necessity, at the same time, doctrinally dispositive. While the limits of legal rules are almost inevitable in most living areas of law, it is most conspicuous in the corporate context in tensions, explained above, between the company's shareholders and its hired directors. The UK's principal company law is, of course, the Companies Act 2006, which – like all legislation – is a product of the political processes and is thus subject to representative democratic criteria and popular accountability.[17] An Act of Parliament, once established, is often a trade-off

13 The most influential paper on this subject is M Jensen and W Meckling, 'Managerial Behaviour, Agency Costs and Ownership Structure' (1976) 3 Journal of Financial Economics 305.

14 See the Companies Act (CA) 2006, ss 170–177. In particular, s 172(1) makes clear that shareholders are now the intended beneficiaries of corporate productivity and enjoy relative priority under varying circumstances. As provided for under Companies Act 2006, ss 260–264, shareholders have the right to initiate proceedings on behalf of the company for breaches of directors' general fiduciary duties.

15 The shareholders in general meeting retain residual and ultimate decision-making power through statute and common law. For example, under the Companies Act 2006, shareholders have the right to vote on the amendment of the articles of association (ss 33 and 21), approval for all economically significant corporate transactions (ss 177, 182, 190–196), mandatory access to the proxy card (ss 303, 314–317, 338–340), anti-dilution rights (ss 171, 549–551, 561–563), and so forth. What is more, shareholders retain the important entitlement to remove directors 'without cause' under Companies Act 2006, s 168. Shareholders are also formally empowered (concurrently with the board itself) to appoint new directors by way of ordinary resolution under the company's articles, although this default rule is formally subject to variation by individual companies. However, in the case of Premium Listed companies it is reinforced by the express UK Corporate Governance Code requirement – applicable on a 'comply or explain' basis – that directors be (re)elected *by shareholders*, which in larger FTSE 350 companies at least should take place on an annual basis. See Financial Reporting Council, UK Corporate Governance Code (July 2018), Code Provision 3.18 <www.frc.org.uk/getattachment/88bd8c45-50ea-4841-95b0-d2f4f48069a2/2018-UK-Corporate-Governance-Code-FINAL.PDF>.

16 This will be discussed in section 3.

17 On this point, see D Feldman, 'Preface' in D Feldman (ed), *Law in Politics, Politics in Law* (Hart 2014) I.

between, on the one hand, distinct practical and political constraints and, on the other, what the legislature intended. This arrangement inevitably gives rise to the articulation of conservative and/or indeterminate legal rules.[18] Simultaneously, the English courts have a general obligation in construing statutes to make law that effectuates legislative purpose.[19] Where the legislative purpose is extremely vague or the delegation of the law-making power to the courts broad, the courts' interpretative role proceeds as a circumscribed formalistic style of legal reasoning.[20] In this regard, the general obligation of the courts is to interpret and apply rigidly the prescriptions of the legislature, or the generalised principles deduced from a series of precedents, to individual disputes, which must be fitted into the existing body of the law.[21]

Ultimately, the ingrained legislative conservatism and formalistic climate of the common law, taken together, are likely to have a dampening effect on the perceived impact of company law's weighted shareholder presence. The following examples provide merely a flavour of the limits of the common law tradition, and of statute, in this practical and real-world sense. First, when we think about the perennial questions central to corporate theory, it is impossible to avoid the uncertainty in law about the nature of ownership and control in publicly listed companies.[22] Second, perhaps the clearest expression of company law's ambiguity regarding the intended beneficiaries of corporate production is the fact that directors are required under the duty of good faith to act in the honest belief that their action was taken, *simultaneously*, in the best interests of the shareholders and the company itself.[23] Third, this analytical frame of reference could be extended to the directors' general company law duty of care and skill.[24] In spite of receiving close judicial and scholarly attention in the 1990s, and again during the most recent company law reform project, the duty is incomplete in the sense that the limited case law means it is difficult to know what amounts to a careful process and what is to be expected of directors in order to determine what a typical careful director performing the

18 Some of the scholarly writing on the most recent company law reform project seems to chime with this point. See e.g. S Worthington, 'Reforming Directors' Duties' (2001) 64(3) Modern Law Review 439, 458; R Goddard, '"Modernising Company Law": The Government's White Paper' (2003) 66(3) Modern Law Review 402, 423.

19 Since the enactment of the Companies Act 2006, much of UK company law is now the preserve of statute law, and the instances of the courts' statutory construction of extant companies' legislation are voluminous. For general academic support on this point, see e.g. H L A Hart, *The Concept of Law* (2nd edn, Clarendon 1994) 132–133; H R Hahlo, 'Here Lies the Common Law: Rest in Peace' (1967) 30 Modern Law Review 241, 248.

20 M J Whincop, *An Economic and Jurisprudential Genealogy of Corporate Law* (Ashgate 2001) 2. For an informative account of the altogether more policy-oriented jurisprudential traditions of the US courts, see E B Rock, 'Corporate Law Doctrine and the Legacy of American Legal Realism' (2015) 163 University of Pennsylvania Law Review 2019.

21 See H L A Hart, 'American Jurisprudence through English Eyes: The Nightmare and the Noble Dream' (1977) 11 Georgia Law Review 969, 979. See also, T Allan, 'Questions of Legality and Legitimacy: Form and Substance in British Constitutionalism' (2011) 9 International Constitutional Law Journal 155, 185; W Friedman, 'Limits of Judicial Lawmaking and Prospective Overruling' (1966) 29(6) Modern Law Review 593, 595.

22 For an excellent exposition, and critique, of share 'ownership' as a doctrinal basis for the shareholders' role in corporate governance, see P Ireland, 'Company Law and the Myth of Shareholder Ownership' (1999) 62(1) Modern Law Review 32. For earlier noises about the unsatisfactory legal conception of share ownership, see A A Berle, 'Modern Functions of the Corporate System' (1962) 62 Columbia Law Review 433, 448.

23 Companies Act 2006, ss 170 and 172(1). For a detailed and nuanced analysis of the distinction between a shareholder-first idea and the pre-2006 common law articulation of the company's interests, see D Attenborough, 'Misreading the Fiduciary Duty of Good Faith' (2020 forthcoming) 20(1) Journal of Corporate Law Studies.

24 Section 174 of the Companies Act 2006.

role would do.[25] Fourth, the *ex ante* disciplinary options for shareholders, which are exercised in general meetings, tend to be practically limited in the case of publicly listed companies,[26] where enforcement options typically give way *to ex ante* monitoring by outside directors and institutional shareholders.[27] Fifth and finally, the statutory derivative claim procedure, which notionally allows a shareholder to bring a legal action on the company's behalf in order to remedy wrongs done to it, specifies permission/leave barriers that shareholders are unlikely to overcome, thereby creating little incentive to litigate.[28]

3 The relationship between law and regulation

Due to powerful political and economic priorities, the majority of which were driven over several decades by globalisation and a neoliberal order,[29] the dominant trend in world affairs has been the 'disaggregation of power into myriad spheres of authority' to facilitate market-based 'norms, informal rules and regimes.'[30] On this basis, the state retains a residual role in delineating the contours of companies' regulation, but then 'commands' the market into filling in the substantive content of that regulation.[31] It is commonly regarded, mainly from within the field of neoclassical economics,[32] as a more efficient rule-making approach, as against the perceived deficiencies of state-promulgated law,[33] based upon the overarching theory of 'reflexive law'.[34] At its heart, reflexive law seeks to mobilise the integrative capacities of the markets and institutions outside the legal system to control companies' adverse socio-economic behaviour in the marketplace.

25 On the marginal or peripheral role of the directors' duty of care after the 2008 global financial crisis, see M Moore, 'Redressing Risk Oversight Failure in UK and US Listed Companies: Lessons from the Citigroup and RBS Litigation' (2017) 18 European Business Organization Law Review 733. For an earlier example from the US, which fundamentally centres on the limits of the duty of loyalty and the duty of care, see J Seligman, 'The New Corporate Law' (1993) 59(1) Brooklyn Law Review 1.

26 A prime example is the mandatory 'without cause' removal right under Companies Act 2006, s 168. But see, L E Strine, 'The Soviet Constitution Problem in Comparative Corporate Law: Testing the Proposition that European Corporate Law is More Stockholder Focused than US Corporate Law' (2016) 89 Southern California Law Review 1239. See also, D Attenborough, 'The Vacuous Concept of Shareholder Voting Rights' (2013) 14(2) European Business Organization Law Review 147; A R Keay, 'Company Directors Behaving Poorly: Disciplinary Options for Shareholders' [2007] Journal of Business Law 656.

27 M Gelter and G Helleringer, 'Corporate Opportunities in the US and the UK: How Differences in Enforcement Explain Differences in Substantive Fiduciary Duties' in D Gordon Smith and A S Gold (eds), *Research Handbook on Fiduciary Law* (Edward Elgar 2018) 331.

28 On the infrequent use of the regime, see A R Keay, 'Assessing and Rethinking the Statutory Scheme for Derivative Actions under the Companies Act 2006' (2016) 16(1) Journal of Corporate Law Studies 39, 41.

29 For some useful critical works on neoliberalism, see R Plant, *The Neoliberal State* (Oxford University Press 2010); D Harvey, *A Brief History of Neoliberalism* (Oxford University Press 2007); N Chomsky, *Profit Over People: Neoliberalism and Global Order* (Seven Stories Press 1998).

30 J N Rosenau, 'Governing the Ungovernable: The Challenge of a Global Disaggregation of Authority' (2007) 1(1) Regulation and Governance 88, 88. See also, S Picciotto, 'Constitutionalizing Multilevel Governance?' (2008) 6(3) International Journal of Constitutional Law 457; B Lange, 'Regulatory Spaces and Interactions: An Introduction' (2003) 12(4) Social and Legal Studies 411, 413.

31 This definition follows closely the one provided in Kershaw (n 1), Part IV. It can be contrasted with 'self-regulation', which, as Julia Black has remarked, implies no particular relationship with the state. See J Black, 'Constitutionalising Self-Regulation' (1989) 59(1) Modern Law Review 24, 27.

32 See e.g. C Sunstein, 'The Paradoxes of Regulation' (1990) 67 University of Chicago Law Review 408.

33 J Black, 'Paradoxes and Failures: 'New Governance Techniques and the Financial Crisis' (2012) 75(6) Modern Law Review 1037, 1038.

34 The pioneering article is G Teubner, 'Substantive and Reflexive Elements of Modern Law' (1983) 17 Law and Society Review 239. See also, G Teubner, *Law as an Autopoietic System* (Blackwells 1993) at ch 5; R M Unger, 'Legal Analysis as Institutional Imagination' (1996) 59(1) Modern Law Review 1.

There are powerful economic arguments for the application of reflexive law strategies to corporate regulation,[35] which cohere to political theories rooted in autonomy and the promotion of individual freedom of choice – and, importantly, while implicating some kind of normative commitment, do not rely on the imposition of binding standards or on a regime of formal sanctions.[36]

Placed in historical context, the UK's Parliament has consciously and deliberately entrusted certain important rule design and enforcement responsibilities to market-based regulatory regimes.[37] These regulatory forms are customarily regarded as hybrid state/market bodies, which provide scope for flexibility, diversity and opt-out at the point of intra-corporate application.[38] While these alternative mechanism diverge considerably in micro-institutional identities and drivers, these lower order rule-sets reveal two ultimate and coherent functions for society. This common purpose, to a large extent, complements the abovementioned role and function of the higher order rules of company law. The UK's market-based regulation attempts in general to leverage the market's knowledge and information advantages to coordinate and enforce the 'rules of the game' governing the rights, relations and conduct of, and within, listed companies whose shares are admitted to trading on the London Stock Exchange.[39] Second, it is in general designed to protect the integrity of the UK equity markets from market misconduct in order to mitigate the information asymmetry problems and agency costs incurred by shareholders in listed companies.[40] The interests of other socio-economic market participants extraneous to considerations of shareholder welfare are in general of secondary importance.[41] No doubt, it would be inaccurate to suppose that company law does not 'matter' in some way to the background state of legal normality against which many areas of corporate life take place. However, market-based solutions have emerged to help increasingly structure and coordinate productivity organised in companies.

This regulatory impact is most evident where, as explained above, Parliament has deferred law-making competence to the market because it neither has the time, nor the interest, nor the expertise. Yet, it is also the case where there is proactive intervention or a role for law central to the control and enforcement of socially suboptimal behaviour, but which the legal rules in question are articulated in only the barest form. Framing important legal questions and issues around seemingly indeterminate legal doctrines and legal decisions is thus often, and understandably, somewhat limited in effectiveness and

35 On this point, see e.g. Teubner, 'Substantive and Reflexive Elements' (n 34).

36 The classic accounts of Hayek's distinction between constructed legal rules and decentralised law are: F Hayek, Law, *Legislation and Liberty* (Routledge 1973) 72–91; F Hayek, *The Road to Serfdom* (Routledge 1944) 75–90. Compare the work of legal realist scholars who, nearly a century ago, revealed the public legal underpinnings of private economic power. See e.g. M R Cohen, 'Property and Sovereignty' (1927) 13 Cornell Law Review 8.

37 See e.g. A Fox, *History and Heritage: The Social Origins of the British Industrial Relations System* (Allen & Unwin 1985).

38 Moore (n 5) 167.

39 Kershaw (n 1) 869.

40 Lord Alexander of Weedon, 'Judicial Review and City Regulators' (1989) 52 Modern Law Review 640, 647; B R Cheffins, *Company Law: Theory, Structure and Operation* (Oxford University Press 1997) 18.

41 For example, not all concerned parties necessarily have input into the formulation of market-based regulatory products. To some extent, this deficit can be addressed through appropriate consultative processes by standard-setting bodies. However, perhaps because of a common investor-protectionist ethos of these regulatory bodies, transparency or participation for non-members has not typically been a priority. In a similar way, the rights of citizens are indirectly affected by, say, the Takeover Panel's decisions and, arguably, only those owning securities may in a technical sense be said to have assented to the situation. On this second point, see *R v Panel on Takeovers and Mergers, ex parte Datafin* [1987] 2 WLR 699, as per Lord Donaldson MR at 838.

outlook. Importantly, in order to interpret and apply the rules in question, it might be tempting for policymakers or practitioners to draw upon more workable and consistent criteria set forth in market-based regulation. These alternative sources have the potential to be used to identify the functions, and context-specific expectations, of directors when fleshing out what should be or should not be done. For our purposes, therefore, market-based rule production might be used not simply as 'an affair of technical bureaucratic minutiae, the thrust and parry of setting agendas, framing issues, and deciding priorities',[42] but also to 'manage the tensions between the "social" and "economic" goals of modern democracies, tensions that enflame passionate and highly wrought political conflict over the ethical limits of global capitalism'.[43] There is already some evidence of such an approach. Accordingly, what follows is a brief outline of the UK's empirically significant market-based regulatory creations, but only to the extent necessary to explain the substantive area they are regulating and the interaction with the substantive law. More will be said later in the article about the institutional arrangements of market-based forms and how their relationship to the state, and their legitimacy, is calibrated through these arrangements.

3.1 The Corporate Governance Code

The Corporate Governance Code of the Financial Reporting Council (FRC) applies, on a 'comply or explain' basis,[44] to all companies that have their shares publicly quoted in the UK.[45] First introduced in 1992, it goes further than closely reflecting existing boardroom practices and has, for example, shifted incrementally, then radically, board composition and separation practices. In the absence of direct attention by UK company law,[46] the Code has been primarily concerned with providing independent governance recommendations that address the structure and function of the board as a disciplinary device. It addresses, in particular, corporate culture; the division of executive and non-executive directors; board and committee structure and staffing of these committees; remuneration guidance; and internal controls.[47] Although the latest version now provides, for the first time since its inception, regulatory recognition of the importance of non-shareholder corporate stakeholders,[48] a pro-shareholder regulatory residue arguably remains an important stimulus of how the FRC understands its role in the corporate

42 B Morgan, 'The Economization of Politics: Meta-Regulation as a Form of Nonjudicial Legality' (2003) 12(4) Social and Legal Studies 489, 490.

43 Ibid.

44 The Listing Rules and Listing Requirements reinforce the Code's voluntary status. See Listing Rules 9.8.6(5)–9.8.6(6).

45 See UK Corporate Governance Code <www.frc.org.uk/getattachment/88bd8c45-50ea-4841-95b0-d2f4f48069a2/2018-UK-Corporate-Governance-Code-FINAL.pdf>.

46 Although the Companies Act 2006 stipulates the minimum number, and age, of directors in a public or private company, the statute remains silent on internal governance arrangements generally, preferring instead to leave such matters to be decided at the intra-firm level.

47 The Corporate Governance Code 2018, ss 1–5.

48 The 2018 version of the Code essentially takes its lead from the Green Paper consultation on Corporate Governance: BEIS, *Corporate Governance Reform* (November 2016). See e.g. The Corporate Governance Code 2018, 'Introduction'.

governance edifice.[49] Consider, for example, the continuing importance attached to *the views of major shareholders* in respect to governance and performance,[50] the senior non-executive director's role as intermediary for the other directors *and shareholders*,[51] and the various sub-committees' main role and responsibility to provide pertinent information to, or engage with, *shareholders collectively*.[52] Furthermore, while listed companies have no choice but to comply with the Code, they can choose to adopt a different approach if that is more appropriate to their circumstances. Where they do so, however, they are required to explain the reason, through their annual report, to the markets generally and, *in particular, their shareholders*, who assess and respond to non-compliance through the credible threat of 'voice' or 'exit'.[53] The Code thus facilitates agency cost-reducing mechanisms *designed to engender investor empowerment and investor confidence* in the board's activities and contributions *by increasing the accountability of directors to shareholders*.[54] Although nothing in the Code overrides or is intended as an interpretation of directors' duties, its weighted shareholder presence has already influenced the courts' interpretive role in various cases. For example, recent case law has begun to use market-based criteria about a director's function and role in order to understand the substantive expectations generated by the duty of care obligation.[55] Similarly, directors in the UK are customarily focused on shareholder value through higher share prices, and an apparent benefit of complying with the Code is an optimal (or at least relatively high) share price.[56]

3.2 The Listing Rules and Listing Requirements

A company must adhere to the Listing Rules and Listing Requirements of the Financial Conduct Authority (FCA) in order to maintain a full listing on the London Stock Exchange.[57] This financial regulatory device encompasses a set of market-based, but contractually supported, obligations applied to premium listed companies, and is subject to the oversight of the UK Listing Authority. In some sense the primary focus of the Listing Rules on ensuring the integrity of financial markets and capital formation might be viewed as having a public policy effect because an increasing amount of the

49 Of course, since 2014 the Code has instructed boards 'to present a fair, balanced and understandable assessment of the company's position', the so-called 'viability statement', which might be regarded as permitting creditors and suppliers to companies to assure themselves of getting a return on their investment. However, the presumptive recipients of the audit, risk and internal control recommendations are primarily the company's shareholders, a fact that is reflected in the Code's express stipulation that this information, set forth in the annual report, 'is necessary for shareholders to assess the company's position, performance, business model and strategy'. See The Corporate Governance Code 2018, 4.26–27.

50 Ibid 1.3.

51 Ibid,2.12.

52 Ibid e.g. 4.25, 4.27, 5.41.

53 On the relative effectiveness of controlling management in this regard, see e.g. J Parkinson, 'The Role of "Exit" and "Voice" in Corporate Governance' in S Sheikh and W Rees (eds), *Corporate Governance and Corporate Control* (Cavendish 1995) ch 3.

54 The Cadbury Report, *The Financial Aspects of Corporate Governance* (1992) <www.ecgi.org/codes/documents/cadbury.pdf>.

55 Perhaps the most prominent example can be found in *Secretary of State for Trade and Industry v Baker (Re Barings)* [1999] 1 BCLC 433. In the Australian context, see *Australian Securities and Investments Commission v Rich* 44 ACSR 431 (2003); *ASIC v Healey* 278 ALR 618 (2011).

56 A Keay, '"Comply or Explain" in Corporate Governance Codes: In Need of Greater Regulatory Oversight?' (2014) 34(2) Legal Studies 279, 283.

57 See UK Listing Rules <www.handbook.fca.org.uk/handbook/LR.pdf>. For an indepth discussion and analysis of the Listing Rules and Listing Requirements, see generally E Ferran and L Chan Ho, *Principles of Corporate Finance Law* (2nd edn, Oxford University Press 2014) ch 13.

population's non-occupational income provision now relies on equity markets.[58] However, it is today the investment intermediaries, rather than ultimate contributors to equity, which have become putative stewards over corporate governance. This growth of intermediation has led primarily to an increased potential for misaligned incentives and a tendency to view market effectiveness through the eyes of intermediaries.[59] Against this backdrop, the Listing Rules contain a number of central corporate governance rules and the 'comply-or-explain' obligation underlying the Corporate Governance Code is also to be found there. The current rules contain six overarching 'Listing Principles', as well as detailed continuing obligation rules in areas such as regulatory notifications,[60] annual financial reports[61] and corporate governance.[62] The basic control is *ex ante* disclosure to the markets generally *and the company's shareholders specifically*, on which it is incumbent to determine whether the response of the company to the Listing Rules does enough and then take some action if they do not. For example, financial information must be disclosed *to domestic and overseas investors* in order to assist them to make active and properly informed decisions about a company's financial position and ensure the pricing of shares in the market is based on adequate and accurate information. In Listing Rule 10, moreover, *shareholder approval is imposed exogenously on all economically significant transactions.*[63] Because shareholder voting is mandatory and binding, rather than a choice variable of the board or management, it acts as a practically effective lever of board accountability. Ultimately, while the Listing Rules, until relatively recently, had a much more direct impact on the terms of a listed company's constitution,[64] *these rules have, over the years, driven the terms of engagement between a listed company and its shareholders.*[65]

3.3 The Takeover Code

For directors of UK listed, or widely held private, companies that are involved in changes of corporate control, the Takeover Code provides the main domestic regulatory framework.[66] The application of the rules is reinforced mainly by the threat of reputational administrative sanctions for non-compliance, with only limited available

58 On this generally, see M Gelter, 'The Pension System and the Rise of Shareholder Primacy' (2013) 43 Seton Hall Law Review 909, 911. For earlier and similar thinking on this issue, see A A Berle, 'For Whom Corporate Managers are Trustees: A Note' (1932) 45 Harvard Law Review 1365, 1368.

59 J Kay, *The Kay Review of UK Equity Markets and Long Term Decision Making* (2012) Executive Summary, point x <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/253454/bis-12-917-kay-review-of-equity-markets-final-report.pdf>. See also, R J Gilson and J N Gordon, 'The Agency Costs of Agency Capitalism: Activist Investors and the Revaluation of Governance Rights' (2013) 113 Columbia Law Review 863.

60 Listing Rules 9.6 and 15.6.

61 Listing Rule 9.8.

62 Listing Rules (mainly) 7, 9, and 15.

63 But also in the case of all UK-registered companies, see the Companies Act 2006, ss 177, 182 and 190–196.

64 R Nolan, 'The Continuing Evolution of Shareholder Governance' (2006) 65(1) Cambridge Law Journal 92, 112.

65 I MacNeil and A Lau, 'International Corporate Regulation: Listing Rules and Overseas Companies' (2001) 50 International and Comparative Law Quarterly 787. On the continuing importance of Listing Rule 10, see M Becht, A Polo and S Rossi, 'Does Mandatory Shareholder Voting Prevent Bad Acquisitions?' (2016) Review of Financial Studies 29(11) 3035.

66 See the Takeover Code <www.thetakeoverpanel.org.uk/wp-content/uploads/2008/11/code.pdf?v=1Apr2019>.
On the historical origins and early development of the Takeover Code, see D Kershaw, *Principles of Takeover Regulation* (Oxford University Press 2016); A Johnston, *The City Takeover Code* (Oxford University Press 1980).

recourse to the courts.[67] Although the Code has been amended over time since its inception, its central provisions have not been altered in substance. Today's Code is an interventionist and onerous body of six general principles and more specific supplementary rules, which proceed on the basis of ensuring the efficiency and procedural fairness of a bid *from the perspective of the shareholders*.[68] In the wake of Kraft's takeover of Cadbury in 2010,[69] the views of employees now receive some prominence,[70] but such changes in general remain 'a halfway house'.[71] Properly understood, the Code remains 'ambivalent regarding the treatment of non-shareholder stakeholders'.[72] Against this pro-shareholder backdrop, the most important obligations of the Code's provisions include: the requirement of *similar treatment of shareholders* of a particular class,[73] the requirement of full and accurate information *for shareholders*[74] and the avoidance of a false market, which is sought through requiring full disclosure of transactions *to the company's shareholders*.[75] One cannot underestimate the fact, moreover, that *shareholders collectively* are entitled to exercise the final 'say' over the outcome of a contested takeover bid by virtue of the so-called 'board neutrality' rule.[76] This managerial authority-limiting rule is a powerful driver of an open market for corporate control, which is customarily regarded as a core mechanism for aligning directors' interests with those of the company's shareholders.[77] Andrew Johnston observed in this regard that the Code's priorities 'reflect the fact that, historically, the primary concern of the drafters of the Code was *to maintain investor confidence in the City* rather than to improve standards of corporate governance generally [emphasis added]'.[78]

67 L Hilliard, 'The Takeover Panel and the Courts' (1987) 50 Modern Law Review 372, 378, remarking that, '[d]espite the potentially wide ambit of Datafin, clearly in future the courts will not wish to supervise the activities of every decision-making body in the country'. Support for this point is found in an excellent analysis of the law in Black (n 31) 32–43.

68 For an academic commentary, see J Armour and D Skeel, 'Who Writes the Rules for Hostile Takeovers, and Why?' The Peculiar Divergence of US and US Takeover Regulation' (2007) 95 Georgetown Law Journal 1727, 1730.

69 B Morris, 'The Cadbury Deal: How it Changed Takeovers' (*BBC News*, 2 May 2014) <www.bbc.co.uk/news/business-27258143>.

70 City Code on Takeovers and Mergers, rules 2.7, 2.9, 20, 24, 25, and 30.

71 W Hutton, C Mayer and P Schneider, 'The Rights and Wrongs of Shareholder Rights' (2017) 40 Seattle University Law Review 375, 395.

72 L Rybak, 'Takeover Regulation and Inclusive Corporate Governance: A Social-Choice Theoretical Analysis' (2010) 10 Journal of Corporate Law Studies 407, 408.

73 City Code on Takeovers and Mergers, rule 11.1 (which fleshes out General Principle 1).

74 Ibid rule 21.3 (which fleshes out General Principle 2).

75 Ibid rule 2.2 (which fleshes out General Principle 4).

76 Ibid rule 21 (which fleshes out General Principle 3). Although company law's proper purpose doctrine, as set out under Companies Act 2006, s 171, provides purposive constraints on post-bid corporate action *primarily intended* to have a defensive effect, it does not, unlike rule 21, prohibit defences that have a defensive impact but which can be characterised as non-defensive business decisions. On this point, see D Kershaw, 'The Illusion of Importance: Reconsidering the UK's Takeover Defence Prohibition' (2007) 56 International and Comparative Law Quarterly 267, 289.

77 See e.g. H Manne, 'Mergers and the Market for Corporate Control' (1965) 73 Journal of Political Economy 110. What is more, the Takeover Code has, in its time, had an impact on the development of the common law. See, for example, *Gething v Kilner* [1972] 1 WLR 337; *R v Panel on Takeovers and Mergers ex.parte Datafin plc* [1987] 2 WLR 699; *R v Panel on Takeovers and Mergers ex.parte Guiness* [1989] 1 Sll ER 509.

78 A Johnston, 'Takeover Regulation: Historical and Theoretical Perspectives on the City Code' (2007) 66(2) Cambridge Law Journal 422, 451. It is important to note, however, that there exists wide disagreement about the nature and drivers of the Code. The literature on the Takeover Code has one strong view that institutional investors captured this regulatory space, while another looks to the influence of investment bankers. See, respectively, Armour and Skeel (n 68); Kershaw (n 66) ch. 3.

3.4 Summary of market-based regulation

On the article's first question about legal-regulatory 'fit', UK company law and market-based rules, rather than being framed as antagonistic alternatives, or mutually exclusive directions of travel, work *together* in a number of core institutions of the UK's company law and governance, albeit with carefully defined, essentially consistent, and mutually complementary roles. These functions are to establish efficient building blocks of productivity organised in companies, understood as looking for rules and structures that maximise shareholder value by reducing the agency costs between directors and shareholders. However, this analysis poses in a conspicuous form the limits or edges of the law's effectiveness and outlook. This includes the context in which market mechanisms address areas of corporate activity that are clearly not covered by company law. Moreover, in other areas where both regulatory systems overlap in relation to their areas of rule production and enforcement, it is evident that the substantive law has sometimes remained curiously indeterminate. On this basis, at least in respect to listed companies, certain basic background rules of company law become little more than a rough 'freehand drawing' onto which the more dynamic and semi-autonomous market-based regulation has the potential to 'paint' definition and colour as reflected in the expected preferences of advisors and their clients. If this is a plausible assessment of the likelihood of regulatory 'spread', then it is arguably clear that its normative and value-laden shareholder presence portends to structure and amplify shareholder interests and value pressures, which is likely to further reduce management insulation from shareholder value pressures.[79] The main problem about market-based regulation in this regard is that it too readily transforms controversial objectives and political issues into questions of rulemaking for hybrid state/market regulators. In this way, it takes critical decision-making powers out of the political process. Since that process is the only way the general population is able to engage, however indirectly, in the shaping of law, this is arguably contentious from a democratic legitimacy perspective.

4 On the political legitimacy question

Following on from the analysis above, the second, and substantive, question to be addressed in this article centres on the idea that the way in which companies are governed, and the way such governance is perceived, contributes centrally to the political legitimacy and democratic accountability of companies' regulation and this underpins both the acceptance and the effects of corporate activity. At any time, an inquiry into the legitimacy and accountability of the regulator constituencies is worthwhile. However, the different ways in which alternative regulatory forms have come, either formally or in practice, to mould and influence the field of company law reveals a more pressing dimension. Accordingly, the question can be put as follows: do these ostensibly independent and apparently politically neutral-technocratic regulatory formations retain a firm basis on which to rest their continuing legitimacy? As this question immediately suggests, a great deal depends on how we understand the concept of law's 'legitimacy'. The organising principles of public law, namely, political and democratic theory, provide us with a useful framework for thinking about questions of legitimacy. This understanding tends to be deployed to explain or justify normatively the institutional legitimacy and deliberative quality of state interventionist functions.[80] Although company

79 H Muir-Watt, 'Private International Law Beyond the Schism' (2011) 2 Transnational Legal Theory 347, 347.

80 There is a long list of works in this tradition; perhaps the most prominent examples include: R Dworkin, *Taking Rights Seriously* (Harvard University Press 1977); R Dworkin, Law's Empire (Harvard University Press 1986). See also, J Raz, *Authority of Law* (Oxford University Press 1979).

law is not customarily characterised as an aspect of 'public' law, this approach is appropriate because a closer inspection reveals that much of the rulebook that affects the UK's system of company law is regulatory or 'non-private' in nature.[81] It is emphasised, by way of a disclaimer, that this section does not seek to flesh out all aspects of the political legitimacy of law, which would not be possible in one article of limited length. At appropriate points the section identifies existing literature that provides further treatment.

4.1 Framing the research enquiry

The first task, then, is to identify what we mean when we talk about democratic criteria or, more aptly, democratic theory, which is an established subfield of political theory. Without too great an over-simplification, the literature is concerned primarily with examining central descriptive questions about how policy or law is made and by whom, but also seeks to realise realistic normative accounts that show us the moderate standards to which we ought to be holding our law-making. Against a backdrop of relatively settled ideas about the rule of law and the separation of powers (especially the independence of the judiciary), explanations and predictions of democratic theory are typically diffused across a variety of academic disciplines and methodological orientations. Correspondingly, there exists broad disagreement about the relative importance of the notion of participation between, and amongst, the theories of earlier writers (the so-called 'classical theorists')[82] and the more recent work (of 'liberal' social scientists).[83] Generally speaking, early democratic theorists subscribe in the main to the view that *the capacity for collective self-governance must be realised through participation*, although one could infer that a delegation of sovereignty by the people to elected officials – as a 'second-best' – is invariably assumed.[84] However, the typical starting point for more contemporary theories of democracy see little scope for *full* participation in the vast complexities of modern governance where technical competence, administrative expertise and executive decisions are needed. When viewed through the influential writing of Joseph Schumpeter, all that is entailed for democracy to work effectively is that *enough* citizens participate to keep the institutional arrangements working satisfactorily.[85] The focus of his theory is, instead, on the minority of leaders, who must be active, initiate and decide, and it is competition between leaders for votes that is the characteristically democratic element in this political method. Whatever the differences between democratic criteria, all of these theories share a basic commitment to 'input-legitimate' governing processes, which are derived from the aggregate pluralistic preferences of the population at large, and that specific 'output-legitimate' institutional arrangements are conducive to policy choices that are generally

81 See generally Moore (n 5); M Stokes, 'Company Law and Legal Theory' in W Twining (ed), *Common Law and Legal Theory* (Blackwell 1986) 155.

82 See the likes of J J Rousseau, *The Social Contract*, M W Cranston (trans) (Penguin 1967); J S Mill, *Representative Government* (Everyman 1910).

83 The most obvious example is J A Schumpeter, *Capitalism, Socialism and Democracy* (Allen & Unwin 1943). An understanding of the nature of Schumpeter's theory is vital for an appreciation of more recent work in democratic theory, in which his established analytical framework and definition of democracy have all become almost universally accepted. More or less an example of this is R Dahl, *After the Revolution? Authority in a Good Society* (Yale University Press 1970).

84 C Pateman, *Participation and Democratic Theory* (Cambridge University Press 1976) 28 (discussing, in particular, the later work of J S Mill).

85 Schumpeter (n 83) 283.

acceptable in terms of consensual notions of the public interest.[86] The concepts of input and output-legitimacy, taken together, constitute the solidaristic prerequisites in debates about what makes substantive company law politically legitimate and democratically accountable.

4.2 The political legitimacy of law

Typically, the UK's traditions of legal and political thought reflect the idea that:

> … [m]uch of what legitimates [substantive] law and distinguishes it from other forms of normativity *are the processes by which it is created and applied – adherence to legal process values, the ability of actors to participate and feel their influence, and the use of legal forms of reasoning* [emphasis added].[87]

Consider, for example, the UK's primary legislation relating to corporate activity, which is, it will be recalled, the Companies Act 2006. As a product of Parliamentary law-making, it derives a substantial part of its input-legitimacy from the formal elements of the rule of law, processes of consensual democracy, and the role of the courts as the ultimate authority on the meaning of legislation. Typically, when we think about how the rule of law supports and nourishes the legitimacy of the English legal system, we look at its restriction of the arbitrary exercise of power by subordinating it to well-defined and established laws that are publicly promulgated, equally enforced and independently adjudicated. The formal and procedural prerequisites of the rule of law are 'designed to ensure the authoritative legal rule is capable of ruling, and that is it is clear, prospective, general, relatively stable and so forth'.[88] On this account, the rule of law ensures corporate life or a substantive review of the corporate framework is ruled by law, order and (in the formal sense) justice – not executive whim, not financial influence and not partisan zealotry. The standard view of the rule of law is, therefore, that it preserves political legitimacy and accountability.

What is more, the UK's political agenda and decision-making process itself must achieve compliance and conformity from the regulated constituencies themselves. Undoubtedly, consensual democracy underpins the consensus-based decision-making process of the design and enforcement of corporate legal rules. As a general rule, reaching consensus, especially as far as different societal goals are concerned, is a characteristic feature of modern democracies. In addition to taking into account as broad a range of opinions as possible, legislatures are democratically elected and accountable bodies. Their members are elected as legislators, and they can be replaced at regular intervals if their constituents dislike what they or their political party are doing in the legislature.[89] Yet the state must not only generate and enforce its rules in a deliberative, impartial and procedurally correct way. From an output-perspective, meanwhile, the state must also ensure that it provides solutions for the society in which it functions.[90] Indeed,

86 This distinction was developed by Fritz Scharpf during the 1970s and is utilised here as it is generally a useful distinction for analysing issues of legitimacy. For a useful synopsis, see F Scharpf, *Governing in Europe: Effective and Democratic?* (Oxford University Press 1999) 6–10.

87 M Finnemore and S J Toope, 'Alternatives to "Legalization": Richer Views of Law and Politics' (2001) 55(3) International Organization 743, 750.

88 J A Grant, 'The Ideals of the Rule of Law' (2017) 37(2) Oxford Journal of Legal Studies 383, 384. Some earlier examples of this point include: J Raz, 'The Rule of Law and its Virtues' (1977) 93 Law Quarterly Review 195, 199–201; L Fuller, *The Morality of Law* (Yale University Press 1964) esp ch 2.

89 See e.g. J Waldron, 'Representative Lawmaking' (2009) 89 Boston University Law Review 335, 335; J Waldron, 'The Core of the Case Against Judicial Review' (2006) 115 Yale Law Journal 1346, 1361.

90 Scharpf (n 86) 10. See also, W Sadurski, 'Law's Legitimacy and "Democracy-Plus"' (2006) 26(2) Oxford Journal of Legal Studies 377, 377.

the legislature in the UK, and elsewhere, has long been required to assess *ex ante* the potential effects of proposed laws through regulatory analysis[91] and wide consultation of the proposed benefits of the legal intervention, as well as parliamentary scrutiny.[92] Eilis Ferran, with a paper published in 2005, opines that:

> … modern theories of good governance emphasise the consultation and collaboration as techniques that can enhance legitimacy of rules and the accountability of the rule-makers. Experience in the UK company law review is certainly supportive of the legitimacy-enhancing function performed by consultation.[93]

In spite of Parliament being the ultimate source of legal authority, there are in practical terms significant limitations on its sovereignty because of a separation of powers.[94] As the final interpreters of companies' legislation, the courts perform the important constitutional function of reconciling the legislative direction with the true meaning of the statute and, hence, its consequences for the resolution of particular company law cases.[95] Of course, the judiciary, though acting in the name of the Crown, is an unelected and formally independent branch of government. While an important element of the separation of powers, it opens up the court, particularly in the context of strike-down powers, to the familiar charge that it is thus inherently democratically illegitimate.[96] Clearly, there are important reasons of political thought and democratic principal that cast doubt on the propriety of giving the courts a company law-making function. Yet, it would be a mistake not to recognise that the court's political legitimacy depends precisely on its *distinctiveness* from political branches, albeit subject to orthodox and fundamentally similar conventions and baselines. Indeed, much of what has been said in the paragraph above is just as true of the common law, in the sense that corporate legal adjudication is framed by an important collection of input-oriented references to legal analogy, legal history, convention, the force of justice and social welfare.[97] In particular, the courts tend to interpret and apply rigidly the prescriptions of the legislature, or the generalised principles deduced from a series of precedents, to individual disputes, which must be fitted into the existing body of the law.[98]

The common law tradition of settling disputes by reference principally to pre-existing corporate legal doctrine and legal decisions has deep and broad appeal within the

91 Petrin (n 2) 537 (and accompanying footnotes).

92 The most recent major company law reform project in the UK took place between 1998–2002 and culminated in the Companies Act 2006. The three main consultation documents were: Modern Company Law for a Competitive Economy: The Strategic Framework (DTI, URN 99/654, 1999); Modern Company Law for a Competitive Economy: Developing the Framework (DTI, URN 00/656, 2000); and Modern Company Law for a Competitive Economy: Completing the Structure (DTI, URN 00/1335, 2000). See also, Company Law, Flexibility and Accountability: A Consultative Document (DTI, URN 04/994, 2004), Annex A, 19.

93 E Ferran, 'Company Law Reform in the UK: A Progress Report' (2005) ECGI Law Working Paper 27/2005, 33 <https://ecgi.global/sites/default/files/working_papers/documents/SSRN-id644203.pdf>.

94 For a detailed historical account of the 'pattern of attraction and repulsion' to the idea of the separation of powers, see M J C Vile, *Constitutionalism and the Separation of Powers* (Oxford University Press 1967) 3ff. For the less strict or rigid separation between the branches than the orthodox account would allow, see A Kavanagh, 'The Constitutional Separation of Powers' in D Dyzanhaus and M Thorburn (eds), *Philosophical Foundations of Constitutional Law* (Oxford University Press 2015) 221.

95 Allan (n 21) 157.

96 Waldron (n 89) 1346; M Tushnet, *Weak Courts, Strong Rights: Judicial Review and Social Welfare Rights in Comparative Constitutional Law* (Princeton University Press 2008) ch 2. Cf R H Fallon, Jr, 'The Core of an Uneasy Case for Judicial Review' (2008) 121 Harvard Law Review 1693.

97 P Winfield, 'Public Policy in the English Common Law' (1929) 42 Harvard Law Review 76, 76.

98 See sources cited in nn 22–23 above (and accompanying text).

tradition of liberal political thought. For H L A Hart, the judge's role is best analogised to the 'delegated rulemaking power [of] an administrative body',[99] In this model, courts have authority to make, or amend, rules for unregulated cases and are instructed to do so with reference to the principles and standards established in the authoritative provisions. Hart further argued that 'legal decision-making does not proceed *in vacuo* but always against a background of a system of relatively well established rules, principles, standards, and values'.[100] Of equal relevance are relatively recent accounts about the democratic foundations of common law reasoning: forensic, transparent, reflective of historical social values, rooted in a concept of individual rights and so forth.[101] Others have sought to show that judge-made law can itself become a valuable channel of political participation and thus to agitate for legal change, especially for those who are marginalised and disempowered in the normal political process.[102] From the output perspective, the legal decisions generated by the courts must cohere with the ongoing expression of public values. This point is well articulated by Lord Nicholls of Birkenhead, who remarked *obiter* that, '[f]or centuries, judges have been charged with the responsibility of keeping this [the common law] abreast of current social conditions'.[103] Accordingly, if a court gets out of sync with the legal and political culture, its pronouncements risk being ignored or overruled by the legislature. So understood, this awareness of the political context is an important component of a court's output-legitimacy in company law cases.

4.3 The political legitimacy of market-based regulation

From the analysis above, it is evident that the UK's company law, which is a product of social organisation, necessarily, therefore, derives authority and political legitimacy from fundamental prerequisites of authoritative adjudication and legislation in addition to those relating to a judge's or legislature's formal entitlement to adjudicate or legislate. These prerequisites essentially rest on public democratic criteria and accountability, such as deliberative, impartial and appropriate process-oriented 'inputs'. Simultaneously, this understanding of legitimacy is instantiated through rule-based 'outputs' that aims to protect both shareholders and, to an extent, the public interest in the effective functioning of wealth-generating companies.[104] On the contrary, market-based regulatory formations typically combine the institutional structure and function of state company and capital markets' rulemaking with the narrow incentives and interests of the most important rule-users in the marketplace. Accordingly, there might be a problem with this article's attempt to inscribe the dominant understanding of legitimacy in regulation. The understanding might not actually work as a general structuring principle for market regulatory bodies because their institutional and administrative workings do not readily or plausibly sit within conventional patterns of the rule of law, consensual democracy and accountability to the courts. No doubt, the quality of outcomes-focused or outcomes-based approaches to the exercise of regulatory power, as we will see, is less contentious because of apparent technocratic or elite expertise, efficiency gains and purposive flexibility. On the input side,

99 H L A Hart, *The Concept of Law* (Clarendon 1967) 132–133.

100 H L A Hart, 'American Jurisprudence through English Eyes: The Nightmare and the Noble Dream' (1977) 11 Georgia Law Review 969, 979.

101 See e.g. Allan (n 21) 185. For a US commentary on this point, see e.g. C E Carpenter, Jr, 'The No-citation Rule for Unpublished Options: Do the Ends of Expediency for Overloaded Appellate Courts Justify the Means of Secrecy?' (1998) 50 Southern California Law Review 235, 248.

102 A Kavanagh, 'Participation and Judicial Review: A Reply to Jeremy Waldron' (2003) 22(5) Law and Philosophy 451.

103 *Re Spectrum Plus Ltd* [2005] 2 AC 680 at [32].

104 See Committee on Company Law Amendment (CMD 6659 1945) at [130].

however, the practical operation and rule control of these alternative regulatory systems takes the risk of being at odds with democratic or policy-based criteria, as well as a substantial body of academic theory and, accordingly, is more likely in practice to undermine and inhibit their perceived authority and legitimacy.

Due to the micro-institutional identities and drivers of the UK's respective regulatory formations, there is not one but many versions of rulemaking bodies, which tend to comprise varying iterations of institutional–functional formality and invariably engage distinct linear points between the public/private poles.[105] Even if the argument at play here does not pin down a regulator archetype, it remains tempting to view the varied regulatory authorities charged with drafting and enforcement powers as having been formed by or being connected to the state. Consider, first, the FCA, which, as the UK's official listing authority, is (and always has been) an independent body created by statute, namely, the Financial Services and Markets Act 2000.[106] The FCA is responsible for writing and enforcing the Listing Rules of the London Stock Exchange, which have the status of delegated legislation. In this regard, the body is accountable to the Treasury and to parliamentary scrutiny. In contrast, the FRC is at the present moment in time formally a non-governmental body that is constituted on a private-sector basis in the form of a company by guarantee. However, public administrative dynamics are inscribed informally insofar as the British government confers statutory authority under the Companies Act 2006 to update the Corporate Governance Code.[107] Simultaneously, its chair and deputy chair are appointed by the Secretary of State for Business, Energy and Industrial Strategy. In spite of the enduring refusal of successive UK governments to displace the perceived prerogative of the FRC to determine the substantive content of the Code, it now seems likely that it will be replaced with an independent statutory regulator with new powers, accountable to Parliament, and named the Audit, Reporting and Governance Authority (ARGA).[108] Finally, the Panel on Takeovers and Mergers is a private body, similar to the FRC, upon which statutory functions in relation to takeover bids have been conferred under the Companies Act 2006.[109] This means that the rules set out in the Code have a statutory basis and apply to takeover bids or merger transactions of public and private companies in the UK. On this basis, the Panel and the Department for Business, Energy and Industrial Strategy (BEIS) coordinate loosely with one another on important issues relating to takeover policy and regulation. Overall, the institutional and functional approach to determining 'public-ness' of the regulator constituencies, conceived as the domain of the state, is ultimately conducive to presenting these regulatory forms as expressly authorised and structured by formal governmental interference or influence in a strictly technical sense.

To many, this institutional and functional analysis as a general structuring principle for market-based regulators may portend the end of the story. Is the body in question part of the government? And is the process, conduct, or decision in question one typically public or discharged by government? In short, most of these regulatory bodies have evolved into agencies that have a more clearly defined public dimension. However, a

105 M Moore and M Petrin, *Corporate Governance: Law, Regulation and Theory* (Palgrave 2017) 22.

106 S 73A. The Financial Services Act 2012 amends the FSMA 2000 to formally create the FCA.

107 Companies Act 2006, part 42. See the Statutory (Amendment of the Companies Act 2006 and Delegation of Functions, etc) Order 2012.

108 See BEIS, 'Audit Regime in the UK to be Transformed with New Regulator' (Press Release, March 2019) <www.gov.uk/government/news/audit-regime-in-the-uk-to-be-transformed-with-new-regulator>.

109 Companies Act 2006, ss 942–965. As amended by the Companies Act 2006 (Amendment of Schedule 2) (No 2) Order 2009.

closer inspection of how these alternative regulatory bodies both organise and operate (relative to the UK's substantive company law-making bodies) raises significant doubt about what it is that gives them their input-legitimacy and in what ways they can be held accountable. Although the state often seeks to achieve its regulatory objectives by conferring a mixture of statutory functions and limited delegated powers whilst simultaneously maintaining formal oversight of the regulation, in doing so it takes the risk that the regulator's practical inner workings and control of the rules may bias the rule detail and rule orientation in favour of primary rule-users who are subject to its regulatory authority.[110] In particular, certain of the most influential rule-making bodies in UK corporate governance, such as the FRC and the Takeover Panel, are controlled by market participants, with a majority of appointed board members from the business and accounting worlds and, at best, only a minority of delegated representatives from other regulators or government.[111] Similarly, when we think about how the Takeover Panel's executive, for example, staffs itself, it becomes difficult to ignore the fact that it comprises limited full-time staff, with mainly secondees drawn from a spread of institutional shareholder, corporate practitioner and other financial institutions within the Panel's regulated community. What is more, charging fees and levies to the firms and market actors who are subject, or have regard to, or benefit from market-based rules and enforcement typically funds these regulatory bodies.[112] By providing the market with such a key role in regulatory control and funding we might have valid public choice concerns that the rules will become tilted toward the financial institutions that are broadly representative of the City establishment's institutional shareholder and associated professional communities.[113] Ultimately, the investigation in this subsection turns on the survival and operational public/private distinction of the regulator constituencies and, once identified, it is tempting to the see the market's practical input into the design and enforcement of various rules and practices as unavoidably weighted towards the interests of (private) business and market participants rather than the (public) formal and procedural prerequisites of the rule of law.

Correspondingly, in spite of the separation of powers being a pervasive feature of the UK's exercise of legal power to form and tailor the rules of corporate life,[114] there are valid reservations about an equivalent system of 'checks and balances' in the realm of market regulation.[115] The continuation of a democratic 'overhead', described above, purportedly extends a form of democratic control from more traditional institutions over the new unelected bodies. Yet, we have seen the attempt to extend democratic controls takes the risk of being undermined by the market's presumptive and actual control of the exercise of regulatory power.[116] Moreover, if we look inside the respective bodies' particular constitutional structures, and the participative procedures, the problem of regulatory capture, and so less accountability, risks feeding through into institutional design. Let us consider the perceived concerns in a little more detail. On the Takeover

110 Kershaw (n 1) 869.

111 For current board membership, see <www.frc.org.uk/about-the-frc/structure-of-the-frc/frc-board/frc-board-members>; <www.thetakeoverpanel.org.uk/structure/panel-membership>.

112 See, for example, <www.frc.org.uk/about-the-frc/funding>; <www.fca.org.uk/firms/fees/how-we-calculate-annual-fees>; <www.thetakeoverpanel.org.uk/the-code/fees-and-charges?>.

113 Davies (n 2) 714.

114 See e.g. E Barendt, 'Is there a UK Constitution?' (1997) 18 Oxford Journal of Legal Studies 137; E Carolan, *The New Separation of Powers: A Theory of the Modern State* (Oxford University Press 2009) 18.

115 See e.g. F Vibert, *The Rise of the Unelected* (Cambridge University Press 2007) ch 4.

116 Of course, many valid arguments suggest that the regulator constituencies operate under the continuing threat of state intervention and influence. On this point, see e.g. Cheffins (n 40) 375. Cf Moore (n 5) 170.

Panel, for example, the Panel, its Executive, and various Committees, subdivide substantive responsibilities for writing, adjudicating and enforcing the Code. The informal *ex ante* rulings of the Hearings Committee can be appealed to the independent Takeover Appeal Board.[117] Similarly, although the FCA is constituted, directed and controlled by its board, comprising executive and non-executive directors, it delegates certain functions/powers to the chief executive officer and/or several committees (e.g. auditing, risk and strategy). Any company may apply to an independent Upper Tribunal for a review of any of the rules made by the FCA or to check whether it came to the correct view of the law.[118] There is, taken together, apparently a latent, but embryonic, separation of powers of, and within, the distinct branches of the Takeover Panel and the FCA. However, the fact remains that almost all of the regulator constituencies, however formally or practically partitioned, tend to be comprised of a relatively homogeneous group of people with industry knowledge, and their careers tend to follow the 'revolving door' process, itself a potential means of capturing the regulator through implicitly biasing the regulators' incentives towards the regulated activity.[119] The concept of capture is problematic because it conveys 'a sense of illegitimate expropriation, performed by one powerful group over others, of the resources we might have thought were provided for public interest goals'.[120] A closer inspection of the practices inhering within the regulator constituencies, therefore, calls into question the perceived validity and rigour of the separation of powers of, and within, the regulator constituencies. In consequence, this understanding might explain why the subsequently produced rules, including board rules, are heavily pro-shareholder.[121]

Finally, a point that has been lurking in the shadows of the analysis above should be brought to light. It is that these alternative regulatory systems generate a reasonably high level of outcome-oriented compliance and conformity of, and within, corporate life, *despite* input-legitimacy and democratic concerns. How might these extensive and rigorous rule-sets function effectively and command the respect of all? Properly understood, the market rule-making bodies that sustain contemporary corporate activity are characterised by a political outlook and rule choice that involve norm formation targeted exclusively at market participants as opposed to as broad a range of individuals as possible. Arguably, as Marc Moore and Martin Petrin concisely note, '[t]his has significant implications in terms of how these bodies understand the proper bounds of their regulatory remit, as well as in relation to the scope of public interest or distributional concerns'.[122] Consider, and compare, for example, the review process of the Company Law Review's legislative reform project with the equivalent means by which market bodies create and apply

117 The Appeal Board's Chairman and Deputy Chairman are appointed by the Master of the Rolls and will usually have held high judicial office. The other members of the Board are appointed by the Chairman or Deputy Chairman and will usually have relevant knowledge and experience of takeovers and the Takeover Code. No person who is or has been a member of the Code Committee of the Takeover Panel may simultaneously or subsequently be a member of the Board. The current membership of the Takeover Appeal Board is available here: <www.thetakeoverappealboard.org.uk/membership/>.

118 FCA, Complaints against the regulators: The Complaints Scheme, 2016 <www.fca.org.uk/publication/corporate/complaints-scheme.pdf>.

119 The pioneering work on capture theory is G Stigler, 'The Theory of Economic Regulation' (1971) 2(1) Bell Journal of Economics and Management Science 1. For a superbly concise account, see T Prosser, 'Theorising Utilities Regulation' (1999) 62(2) Modern Law Review 196, 200–206.

120 L G Baxter, '"Capture" in Financial Regulation: Can We Channel it toward the Common Good?' (2011) 21(1) Cornell Journal of Law and Public Policy 175, 176. For a useful account of regulatory capture theory, see Prosser (n 119) 200–206.

121 Davies (n 2) 714–715.

122 Moore and Petrin (n 105) 22.

regulatory solutions. The Company Law Review consulted extensively and was subject to rigorous parliamentary scrutiny on its self-identified 'scope' of company law, however ultimately channelled into the deeper shareholder rights bias in the UK's company law. Yet, the formal consultation process of the Takeover Panel's Code Committee, for example, introduces or amends any rules of the Code through comparatively marginal, if useful, dialogue mainly (although by no means always) with all relevant 'deal insiders'.[123] In spite of making available all responses for public inspection, the presumption and expectation that market participants should address and resolve the problems of their own making, although producing output-legitimacy in a strictly limited sense, risks being less openly and effectively questioned by the general public and in the public policy arena.[124] This is in spite of the fact that the Panel, conceived as functionally silent or neutral in respect to distributional concerns extraneous to considerations of shareholder welfare,[125] is nevertheless likely to, and does, produce significant and extensive socially normative outcomes.[126]

4.4 From political to technocratic legitimacy

Based on the above analysis, it might seem misplaced to concentrate upon dominant accounts of modern democratic criteria, principally because the operation and outlook of market regulatory bodies lie somewhat outside standard legitimacy mechanisms. An apparently weak form of input-legitimacy, and a weighted outcome-legitimacy, relatively speaking, may scarcely come as a new or exciting truth to proponents of market-based regulatory regimes.[127] In neoliberal discourse, political legitimacy and democratic accountability are relatively important prerequisites for regulator constituencies. However, the apparent availability of a greater degree of professional expertise and market-determinable practices (relative to standard public law-making bodies) are often claimed to invoke a rival conception of legitimacy, which supports and nourishes both the general acceptability and the actual effects of regulation in two distinct ways.[128] First, there is the apparently technocratic or practical criteria that generates regulatory solutions in the UK, whereby predominantly financial and legal intermediaries design and enforce rules in an apparently politically neutral context,[129] which is shaped by a set of professional ideals of what regulatory value judgements are economically efficient from the perspective of the marketplace.[130] The concept of technocracy maintains the perceived apolitical prerogatives of the regulator, which then justifies the public trust that is placed in them to exercise significant and extensive regulatory power on a non-arbitrary basis outside of the standard 'checks and balances' of the democratic state framework. Second, these regulatory formations are reputed to closely reflect the endogenous preferences of private business and market participants generally rather than having any social-

123 For a list of the respondents who submitted responses on a non-confidential basis to arguably the most significant Takeover Panel review in the previous decade, see the Takeover Panel, *Review of Certain Aspects of the Regulation of Takeover Bids* (June 2010) <www.thetakeoverpanel.org.uk/wp-content/uploads/2008/11/2010-22.pdf>.

124 Prosser (n 119) 199. See also Moore (n 5) 170–172.

125 Moore (n 5) 172.

126 See *Datafin* (n 41) as per Lord Donaldson MR at 838.

127 Black (n 33) 1042.

128 The following two points on the general acceptability of market bodies follow closely the concise analysis in Moore (n 5) 170–174. See also, Prosser (n 119) 197–199.

129 P Mair, *Ruling the Void: The Hollowing Out of Western Democracy* (Verso 2016) 6.

130 Morgan (n 42) 507. See also, B Gilley, 'Technocracy and democracy as spheres of justice in public policy' (2017) 50(9) Policy Sciences 9, 14–16.

distributional implications in their own right. Accordingly, the respective rule systems could be said to derive their customarily regarded regulatory legitimacy from market compliance and conformity, insofar as they commonly seek to reflect rather than displace the expected economic interests of the advisors and of their clients generally.

However, it is important to sound a note of caution against these two reputed sources of the continuing survival and legitimacy of regulator constituencies in the UK. First, the regulatory *ideal*, that is a standard (or, perhaps, set of standards) to which private or professional arrangements of market forms are said to conform, does not relieve us from the need to look at the evidence for this claim. Clearly, the existence of empirically dispositive evidence to support this account about optimal market solutions in corporate and financial markets regulation is deeply ambivalent.[131] Equally clearly, descriptive accounts of economic efficiency in this regard tend to be conflated with normative values, or are the product of unanalysed assumption and borrowing and the cascading of such borrowings. Rather, the limits of markets as regulatory mechanisms for constraining socially suboptimal behaviour and outcomes are numerous and well documented. On this basis, the most one could say is that 'although a perfect market is superior to an imperfect legal rule, an imperfect market *may be either better or worse* than an imperfect legal rule [emphasis added]'.[132] It follows that in some, but far from all, areas of corporate activity, market-based regulation *might* prove to be the most economically efficient mode of regulation, while in others the substantive law *might* be relatively more effective in this regard. In addition to a questionable inscription of economically efficient ideals in technocratic rule formation, earlier discussion, meanwhile, has queried whether the market-based regulator's claim to political neutrality or impartial pragmatism is necessarily valid. No doubt, while neutrality of method might be difficult to achieve, the regulator constituencies tend to maintain the position that partiality is possible to avoid: technocrats avoid value claims, reject political or ideological predispositions, and often receive general acceptability in the marketplace. One problem with this approach is that expert knowledge is not necessarily neutral, especially when we have observed that regulation may be captured by powerful vested interests, or at least, in light of a weighted shareholder presence in existing regulatory solutions, is not always used for neutral purposes. Ultimately, therefore, one needs to be wary of such broad-brush claims about the plausible, although speculative, legitimacy of the regulator through apparently neutral-technocratic and pragmatic expertise.

The second claim about legitimacy of market regulatory formations refers, it will be recalled, to the single goal of protecting and advancing economic efficiency (rather than social welfare-enhancing regulatory solutions) in the marketplace, which, in turn, generates compliance and conformity, insofar as the respective rule systems purport to crystallise the expected economic interests of industry and market participants generally. Certainly, the constituent rules and practices of, say, the Takeover Code or the Corporate Governance Code often, and understandably, reflect the commercial or political interests of financial and legal intermediaries that might otherwise emerge through private ordering, were the costs of making adequate provision for all possible contingencies sufficiently low. Simultaneously, targeted norm formation also means that some of the most important market-based rules apply, either formally or in practice, only to a small, but important, subset of the total population of companies regulated under the Companies Act 2006 and earlier legislation, i.e. those which are listed on the Main Market

131 J Armour, 'An Economic and Jurisprudential Genealogy of Corporate Law' (2002) 61(2) Cambridge Law Journal 467, 468.

132 M A Eisenberg, 'Bad Arguments in Corporate Law' (1990) 78 Georgetown Law Journal 1551, 1551–1552.

of the London Stock Exchange, and often only to those which have chosen a premium listing. In some sense, if one were to restrict the analytical lens to the exclusive mission goal and targeted audience, explained above, there are valid arguments in favour of describing the UK's approach to company and financial market regulation as legitimate, *provided that the system applies only to the areas of regulation and does not have any wider socially determinative effect in its own right*. Yet, the Takeover Panel or the FCA, for example, in remaining deliberately and consistently silent when it comes to surrounding public policy debate about the attraction or otherwise of their overall social effect, while simultaneously providing for a formal and enduring weighted shareholder presence in corporate governance, could nonetheless have one unintended consequence. Specifying, it is likely that this task-driven method lends significant tacit political or ideological support to the idea that protecting and empowering shareholders is inherently democratic, although not chosen or necessarily supported by the public at large, or even open to meaningful discussion among them.

5 Conclusion

At the beginning of this article, two objectives were specified: the first concerned the relationship and interaction between the UK's substantive company law and its market-based regulation; the second explored the legitimacy of market-based regulation in light of its potential to mould and influence the content of legal rules where the state is passive or the state remains active, but the law is indeterminate. As to the first, although the UK's company law provides the legal ecology to facilitate productivity organised in companies by reducing agency costs between participants, market-based regulators, meanwhile, have become increasingly important to the design and enforcement of basic 'rules of the game' between market participants by reducing socially suboptimal behaviour and outcomes. This is assuredly the case where the British state has deferred law-making competence to the market because it neither has the time, nor the interest, nor the expertise. However, where there *is* proactive intervention or a role for law, the subsequent legal rules are often, and understandably, viewed as somewhat limited in effectiveness and outlook. This is in turn likely to have a number of direct or practical implications, not least the dampening effect on the perceived impact of company law's weighted shareholder presence. On this basis, in order to interpret and apply the rules in question, policymakers or practitioners might be tempted to draw upon heavily pro-shareholder, if more workable, guidance provided in market-based regulatory instruments. There is already some evidence of such an approach. As to the second inquiry about legitimacy, two points can be made. One is that the main problem about market-based regulation in this regard is that it too readily transforms controversial objectives and political issues into the question of rulemaking for hybrid state/market regulators. In this way, it takes critical decision-making powers out of the public process. Since that process is the only way the general population is able to engage, however indirectly, in the shaping of law, this is arguably contentious from a democratic legitimacy perspective. The other is that these market-based regulators do not readily or plausibly lend themselves to dominant political narratives and democratic criteria, which customarily nourish and support procedural and substantive legitimacy of state interventionist techniques in company law-making. Simultaneously, the deployment of a rival conception of legitimacy, conceived as neutral technocratic expertise and market determinable practices to support both the general acceptability and the actual effects of regulation, is arguably contentious because of public choice concerns about regulatory capture and the fact that the general

acceptability or otherwise of market-based rules is typically limited to the marketplace, despite the wider socially determinative impact on the population at large.

This subject of enquiry is timely and relevant for two reasons. First, while this article focuses upon the UK corporate regulatory framework and potential challenges, questions about the relationship between regulatory instruments and techniques, and the relative legitimacy of law-making in company law and financial regulation, are ones that are of general interest to policymakers around the world operating at national, regional and international levels. The internationally mobile character of much modern business operations make it important for states and regions to benchmark their 'mix' against that chosen by others and to ask whether they have achieved a combination that is commercially attractive and likely to promote the contribution of the corporate sector to the overall economic and social well-being of society. Second, the article's findings have significant ramifications for continuing international debates on the devolution of specific functions to market-based regimes accompanied by new types of public–private interactions. Many open questions remain, theoretically and practically, about these changes and their implications for the legitimacy of governance processes. The present article, accordingly, develops this important strand of competing thought by framing the political and constitutional case that market-based regulation risks undermining the rule of law and its democratic potential, with the effect of exacerbating corporate threats to public interests. In doing so, it moves forward the democratic and political case for representative law-making in company law beyond the contemporary neoliberal rationales centred on the relative value of market competence.

NILQ 70(4): 407–423

Theses on the relationship between rights and social struggle

Dr Dimitrios Kivotidis

University of East London

> There is therefore an antinomy, of right against right, both equally bearing the seal of the law of exchange. Between equal rights, force decides.[1]

Abstract

This article examines the relationship between rights and social struggle. This topic is revisited in light of the phenomenon of rising inequality in the aftermath of the last capitalist crisis, which reignited the debate on the role of rights in processes of social mobilisation. In this context, this paper examines three very recent contributions to this debate, namely Samuel Moyn's Not Enough: Human Rights in an Unequal World*, Radha D'Souza's* What's Wrong with Rights*, and Paul O'Connell's work on a critique of the displacement thesis. In critically discussing these contributions it introduces and elaborates on six theses which describe the relationship between rights and social struggle. The argument focuses on the important role of rights in the struggle between different social forces, as well as their limitations in promoting a critique of the structural roots of social inequality.*

Keywords: rights; social struggle; inequality; social justice; Marxism

Introduction

Where we now have a right, we used to have the painful void of its absence. But what happens when even constitutionally enshrined rights are rendered meaningless by bourgeois state and international institutions? The question of rights and their place and role in capitalism is not new and has had a constant and polymorphic presence in the legal and political discourses over the last few centuries. Rights and the struggle for rights have been central to the development and formation of bourgeois societies, capitalist social relations and the juridico-political apparatus accompanying these. Meanwhile, rights and struggle for rights have been central to the struggle of the toiling classes, namely the struggle of the working class and popular strata for the improvement of their living and working conditions and the amelioration of conditions of repression and exploitation. Last, but not least, rights in the form of concessions to these toiling classes have been central to the processes of ensuring the reproduction of the bourgeois state and rule, as well as of the capitalist property and productive relations, in the face of intensified contradictions that could have led to revolutionary upheaval and change.

1 Karl Marx, *Capital: Critique of Political Economy* vol I (Penguin Classics 1990) 344.

Therefore, it is safe to claim that the role of rights in capitalism is contradictory. This is due to the fundamental antinomy between 'equal rights' in capitalist society, first identified by Karl Marx. The rights of workers and the rights of capitalists clash. Rights shape and are themselves shaped by class struggle. This is the point of departure in this paper. In Europe, during the last – until the next – crisis of capitalism, we witnessed how constitutional guarantees of social and economic rights turned into empty words on paper. This was achieved either through international agreements (Memorandums of Understanding) for structural adjustment programmes which accompanied the financial assistance towards countries such as Greece, Ireland, Portugal and Spain, or through ordinary legislation which restricted or eliminated the right to collective bargaining and the protection of workers against collective dismissals, as well as limiting the right to strike in countries like France, Italy and Britain; but always in the name of restoring growth and competitiveness of European economies and the economy as a whole. In fact, the elimination or circumvention of social and economic rights was a necessary response to the crisis on behalf of capital, so as to intensify the exploitation of the labour force and, therefore, its own profitability.

Capital throughout the world has had to increase the exploitation of labour and redistribute wealth in its favour, in order to reproduce the conditions of its existence and dominance. In parallel, and as a result of this, inequality rises globally. The process of pauperisation of the majority of the population finds its double in the process of wealth accumulation by very few. According to Credit Suisse, 1 per cent of the global population owns 47 per cent of global wealth.[2] One could safely argue then that not everyone was adversely affected by the economic crisis. The tendency of capital accumulation following the crisis is confirmed in Oxfam's latest report on social and income inequality. According to this report, just eight billionaires now hold the same wealth as the 3.6 billion people who form the poorest half of the world's population,[3] whereas more than 80 per cent of the new global wealth goes to the top 1 per cent while the poorest half get nothing.[4] As a confirmation of this tendency, it is worth noting that, according to Oxfam, in 2016 the number of billionaires holding an equal amount of wealth as the bottom half of the world was 62; 80 in 2014; and 388 in 2010.[5]

In this context, the debate on the relationship between rights, social struggle and social and economic inequality has been reignited. This debate, which is constantly revisited, revolves around the role of discourses and practices of rights in emancipatory processes and social struggles. Human rights practices and discourses have been criticised for legitimising and perpetuating greater injustice than they addressed, as well as for being atomistic and therefore inappropriate for social struggle around common goals. At the same time, there have been various attempts at reconsidering the role of rights and opening them up to a serious evaluation 'by the Left in order to develop a positive

2 Credit Suisse Research Institute, *Global Wealth Report 2018* (Credit Suisse October 2018) 9.

3 Oxfam, 'Eight People Own Same Wealth as Half the World' (Oxfam Press Release, 22 January 2018) <www.oxfam.org.uk/media-centre/press-releases/2017/01/eight-people-own-same-wealth-as-half-the-world>.

4 Oxfam, 'More than 80 per cent of New Global Wealth Goes to Top 1 per cent while Poorest Half Get Nothing' (Oxfam Press Release, 22 January 2018) <www.oxfam.org.uk/media-centre/press-releases/2018/01/more-than-80-per-cent-of-new-global-wealth-goes-to-top-1-per-cent-while-poorest-half-get-nothing>.

5 Oxfam, 'An Economy for the 1%: How Privilege and Power in the Economy Drive Extreme Inequality and how this Can Be Stopped' (Oxfam Policy Paper, 18 January 2016) <www.oxfam.org/en/research/economy-1>.

program for a politics of rights'.[6] According to this approach, rights can and should be deployed in emancipatory political projects because the recognition of rights, such as welfare or socio-economic rights, can form part 'of a series of measures that constitute a dialectics of subversion of the logic of capital'.[7]

This article, rather than reiterate the above positions, will examine them in light of new developments, that is the effect of the recent crisis and the attack on social and economic rights. It, therefore, focuses on three very recent contributions to this debate: Radha D'Souza's *What's Wrong with Rights*,[8] Samuel Moyn's *Not Enough: Human Rights in an Unequal World*[9] and Paul O'Connell's work on a critique of the displacement thesis.[10] In critically assessing these contributions, the article will propose and examine six theses on the relationship between rights and social struggle. These will be introduced in the first section of the article, which sets the theoretical context of the discussion by outlining the central aspects of the critical theoretical approach to rights and human rights. The next two sections will discuss two of the more recent contributions to this debate (i.e. those of Moyn and D'Souza) and will set the tone for the final section, which will further the examination of the role and limits of rights in the epoch of imperialism, by focusing in particular on the right to strike and examining how this specific right has been treated by the bourgeois juridico-political apparatus, as well as the reasons for this treatment.

Theoretical context

There are many sides and approaches to the relationship between rights and social struggle. Correspondingly, the literature on the variety of issues relating to this topic is vast. This paper chooses to focus on the issue of rights exclusively from a Marxist perspective and assess it through a lens that puts emphasis on the issue of social antagonisms and the role of rights in mediating and superseding these. However, a Marxist analysis of rights is situated in the context of the existing critical literature of rights, which is vast and growing. The main parameters of this critique are set by what has been termed the 'displacement thesis'.[11] Despite this being one specific strand of a more generalised critique and dismissal of human rights,[12] it has been argued that this thesis has been the common ground for the critique of rights carried out – for the most part – by the critical legal studies (CLS) movement,[13] in two phases (one during the

6 Amy Bartholomew and Alan Hunt, 'What's Wrong with Rights?' (1991) 9 Law and Inequality 1.

7 See Prabhat Patnaik, 'A Left Approach to Development' (2010) 45(30) Economic and Political Weekly 33.

8 Radha D'Souza, *What's Wrong with Rights* (Pluto Press 2018).

9 Samuel Moyn, *Not Enough: Human Rights in an Unequal World* (Harvard University Press 2018).

10 Paul O'Connell, 'Human Rights: Contesting the Displacement Thesis' (2018) 69(1) Northern Ireland Legal Quarterly 19. Also, Paul O'Connell, 'On the Human Rights Question' (2017) 40(4) Human Rights Quarterly 962–988.

11 O'Connell, 'Human Rights: Contesting the Displacement Thesis' (n 10).

12 See, for instance, Slavoj Zizek, 'Against Human Rights' (2005) 34 New Left Review 115; and Jarret Zigon, 'Human Rights as Moral Progress? A Critique' (2013) 28 Cultural Anthropology 716.

13 For an overview of the CLS movement see: Roberto Mangabeira Unger, *The Critical Legal Studies Movement* (Verso 2015); Alan Hunt, 'The Theory of Critical Legal Studies' (1986) 6 Oxford Journal of Legal Studies 1; and Costas Douzinas and Adam Geary, 'News from Nowhere: Anxiety, Critical Legal Studies and Critical "Tradition(s)" in Costas Douzinas and Adam Geary, *Critical Jurisprudence: The Political Philosophy of Justice* (Hart 2005).

1980s[14] and the second following 9/11 and the US-led imperialist interventions which led to a new wave of human rights critiques)[15] and which remains highly influential today.[16]

This thesis maintains that reliance on the language of 'human' rights by movements for radical social change is problematic, because it tends to crowd out (or displace) other, potentially emancipatory, languages and, as a consequence, distract attention from broader, structural causes of injustice and oppression. Variations of this argument have been put forward by thinkers such as Morton Horwitz, according to whom 'framing issues of social justice in terms of individual rights has the additional effect of denying equal legitimacy to claims that the overall social distribution of wealth and power is unjust',[17] Wendy Brown, who puts emphasis on the difficulty of trying to engage simultaneously in human rights projects as well as all other kinds of 'justice projects',[18] and Austin Sarat and Thomas Kearns, who argued that 'reliance on rights in political struggles and by political movements invites a kind of legal imperialism, in which courts and lawyers take on an unhealthy prominence'.[19]

A comprehensive critique of the displacement thesis has been carried out recently by O'Connell. O'Connell disagrees with this nihilistic approach towards 'human' rights. He argues that the mobilisation of rights language can make an important contribution to movements for radical social change, without displacing or precluding the mobilisation of other emancipatory languages and the challenging of deeper, structural causes of injustice. In support of his argument, he uses the examples of the Focus E15 campaign and the Right2Water movement in Ireland, which 'engaged the language of human rights alongside other frames of reference, and connected the specific rights struggle its protagonists were engaged in with broader causes of injustice'.[20]

This article intends to build on O'Connell's critique of the displacement thesis, by putting emphasis on the need to examine the role of rights in the context of capitalist contradictions and the issue of class exploitation in particular.[21] In this context it seeks to contribute to the ongoing debate on rights and social struggle by introducing and examining the following theses:

14 See Anthony Chase, 'The Left on Rights: An Introduction' (1984) 62 Texas Law Review 1541; and Morton Horwitz, 'Rights' (1988) 23 Harvard Civil Rights–Civil Liberties Law Review 393, 400.

15 See Duncan Kennedy, 'The Critique of Rights in Critical Legal Studies' in Wendy Brown and Janet Halley (eds), *Left Legalism/Left Critique* (Duke University Press 2002) 178; and Wendy Brown, '"The Most We Can Hope for …": Human Rights and the Politics of Fatalism' (2004) 103 South Atlantic Quarterly 451, 453 .

16 See Robin West, 'Tragic Rights: The Rights Critique in the Age of Obama' (2011) 53 William and Mary Law Review 713, 716; and Ratna Kapur, *Gender, Alterity and Human Rights: Freedom in a Fishbowl* (Edward Elgar 2018).

17 Horwitz (n 14) 400.

18 Brown (n 15) 460.

19 Austin Sarat and Thomas Kearns, 'Editorial Introduction' in Austin Sarat and Thomas Kearns (eds), *Identities, Politics and Rights* (University of Michigan Press 1997) 1, 4–5.

20 O'Connell, 'Human Rights: Contesting the Displacement Thesis' (n 10) 27–32.

21 It has to be noted here that the centrality of class struggle in Marxist analyses does not discount or discard the other forms of oppression, be it gender, race, caste or sexuality, but rather encourages a comprehensive understanding of the interwoven nature of these processes. For instance, racial division is ingrained in the logic of capitalism and cuts through the history of capitalist society. One cannot comprehensively make sense of racial discrimination except on the basis of an understanding of capitalist contradiction and Marx's insights on the processes of social reproduction of capitalism. This refers not just to the process of primitive accumulation but also to the *divide et impera* policies of the capitalist to ensure maximum exploitation of the labour force. See also David Roediger, *Class, Race and Marxism* (Verso 2017).

1. 'Human' rights can be recuperated by social agents other than those they seem to be designed for.
2. 'Human' rights can be used crudely to legitimise imperialist interventions.
3. 'Human' rights can be used to divert social struggles away from political goals and canalise them into legal proceedings.
4. The struggle over rights is one form of class struggle.
5. The protection, restriction or elimination of rights may serve either the immediate or the strategic interests of social classes.
6. The struggle of the working class and the popular strata for rights may only serve their strategic interest (i.e. the elimination of the conditions of exploitation) if it is ultimately directed towards the elimination of the capitalist social relations.

The following sections elaborate on these theses, but for the moment it is important to note the distinction they introduce between rights – to which the last three theses refer – and 'human' rights – referred to in the first three theses. This approach follows D'Souza's point that rights should not be reduced to 'human' rights, because the prefix 'human' is a post-Second World War addition 'that conceals what is entailed in rights in the epoch of imperialism'.[22] An extensive analysis of this distinction and its implications could potentially form the subject of a different article, but does not fall completely outside the scope of our argument. For the purposes of this article it suffices to locate the problematic nature of this discourse in the individualising function of the concept of 'human', in its bourgeois conception, as well as the de-politicising effect of the discourse which is discussed later. This argument is based on the well-established critique of the bourgeois conception of 'humanity' which is constructed in the model of the profit-seeking rational actor.[23] The concept of rights, on the other hand, may refer to a broader set of claims.

Contrary to 'human' rights, rights can be associated with the idea of right-qua-justice and can be pursued as collective demands through collective struggle and a comprehensive critique of the totality of social relations of the existing system. The argument developed here is that the rights discourse can play a significant part in emancipatory processes, as long as it is expanded to assume a materialist content; that is as long as it becomes an integral part of a comprehensive critique of the regime of power, property and productive relations in capitalism. In this manner its role could be crucial for uniting the struggle for everyday matters, such as working and living conditions, with the struggle for a different social arrangement altogether that prioritises social needs instead of profit. This unity of economic and political class struggle may cancel in practice the separation between the economic and the political sphere in capitalism, which is central to the reproduction of capitalist social relations.

Moyn's critique of inequality

Let us begin the review of the more recent scholarship on the relationship between rights and social struggle with Moyn's *Not Enough: Human Rights in an Unequal World*[24] Moyn's contribution to this debate is significant because he shifts the focus of the debate towards

22 See D'Souza (n 8) 3, 65, 66.

23 See C B Macpherson, *The Political Theory of Possessive Individualism: Hobbes to Locke* (Oxford University Press 2010).

24 Moyn (n 9).

political economy and the relations of distribution in capitalism to contextualise his argument on the role of rights in debates and practices against social inequality. As a result, Moyn contributes to the opening of an intellectual pathway in mainstream discussions about rights where the discussion and critique of capitalist relations of distribution is considered viable and legitimate – fulfilling thus a function analogous to the work of Thomas Piketty on capital in the twenty-first century in the field of economics. However, as will become evident in our analysis, Moyn's critique of capitalist relations with regards to rights suffers from the same limitation of focus as Piketty's work.

Moyn describes his work as 'an intellectual and ideological history' of human rights.[25] In his critical analysis, human rights are set against the absolute goal of 'full-fledged distributive justice'.[26] He recognises that the age of human rights 'has not been kind to full-fledged distributive justice, because it is also an age of the victory of the rich'. The age of human rights has focused on just one of the two imperatives of distributional justice: sufficiency; paying little to no attention to its counter-part, namely, equality. These two imperatives are different as to their subjects (sufficiency concerns itself with the status of the poor, whereas equality with the status of the rich) and goals (sufficiency concerns itself with providing a 'floor' for the poor, whereas equality with setting a 'ceiling' for the rich). Their difference is what causes tension between the different imperatives of distributional justice.

According to Moyn, the two imperatives have to be pursued simultaneously in order for the ideal of distributional justice to be reflected in a socio-political system: 'Not merely a floor of protection against insufficiency is required, but also a ceiling on inequality, or even a commitment to a universal middle class.'[27] Moyn carried out his intellectual history of human rights 'out of dissatisfaction with mere sufficiency and committed to a more ambitious equality'.[28] On this basis, he considers that human rights can have a central role in this project of promotion of the principle of fair distribution and pursuit of the dual imperative of distributional justice. The reason behind this conviction is his belief that there is no necessary connection between human rights and market fundamentalism.[29] Their connection has been historically contingent, and there is no reason why human rights cannot become vessels for the pursuit of distributional demands of sufficiency and equality.

This conviction conditions Moyn's critical stance with regards to critics arguing for a structural relationship between human rights and neoliberalism, such as Naomi Klein, according to whom 'the human rights movement helped the Chicago School ideology to escape from its first bloody laboratory [in Chile] virtually unscathed'.[30] For Moyn, 'human rights surely did not bring the neoliberal age about'.[31] On the contrary, human rights had a positive impact in citizens' lives by enabling 'unprecedented scrutiny' on state violence and protection of equality in matters of gender, religion, race, or sexual orientation. In addition to this, according to Moyn, there is no reason to think that human rights cannot coexist with more 'structural' politics and critiques of the current social relations. Strangely enough, this point is followed in that very same paragraph with a tentative yet

25 Ibid 10.

26 Ibid 2.

27 Ibid 4.

28 Ibid 10.

29 Ibid: 'There is no reason for human rights ideals to continue the accommodating relationship they have had with market fundamentalism and unequal outcomes.'

30 Naomi Klein, *The Shock Doctrine: The Rise of Disaster Capitalism* (Metropolitan Books 2007) 118–128.

31 Moyn (n 9) 175.

aphoristic defence of neoliberalism which, comparable to Chinese marketisation, 'brought more human beings out of poverty than any other force has in history'.[32]

Nevertheless, Moyn's argument consists of an unconditional acceptance of human rights and, in particular, certain social and economic rights, which he considers part of the general category of 'human rights', to the extent that they promote the twin imperatives of sufficiency and equality. Rights, such as the right to work, the right of subsistence, or the right to the product of one's labour, have to be prioritised in order to promote the goal of distributional justice. At this point, Moyn makes a passing yet highly critical reference to Karl Marx who 'did not embrace distributional equality before the revolution', neither did he 'envision material fairness in a communist state'.[33] According to Moyn, Marxist calls for revolution demanded an end to hierarchical power rather than fairer distribution.[34]

Precisely this point is proof of Moyn's failure to grasp Marx's critique of the capitalist mode of production, as well as the basis for understanding the limits of his argument and his critical analysis of human rights. For, it is common knowledge to those introduced to the Marxist critique of capitalism that not only does this critique move beyond the relations of domination and suppression that structure the problem of power hierarchy, but it also examines distributional relations in capitalism as part of the unified whole they comprise with the relations of production. For Marx, 'any distribution whatever of the means of consumption is only a consequence of the distribution of the conditions of production themselves'.[35] The distributional relations in capitalism are themselves conditioned by the distribution of the means of production; that is by the capitalist relations of production.[36] The distribution of the means of consumption in capitalism results automatically from the 'fact that the material conditions of production are in the hands of non-workers in the form of property in capital and land, while the masses are only owners of the personal condition of production, of labour power'.[37]

Two points follow necessarily from the above. First, Moyn is mistaken in considering the absence of a demand for 'material fairness' as a gap in the Marxist critique of capitalist relations. The only reason Marx and Marxists do not focus on a critique of the distributional relations of capitalism is because this critique forms part of the more comprehensive critique of the totality of capitalist relations of production. The second point, which is directly related to the first, has to do with the concrete limitations of Moyn's analysis of inequality in capitalist society and the role of rights in addressing this. I argue that Moyn fails to assess the structural roots of social inequality in capitalism by focusing solely on demands of distributional justice, while neglecting the issue of exploitation at the point of production. As O'Connell puts it in his review of Moyn's book, inequality in capitalist society is not the result of poor distributional choices, but first and foremost a result of the structural character of the extant social system.[38] Therefore, the fundamental question for addressing the issue of inequality via rights is not how we can use rights to minimise inequality by altering the distributional model and

32 Ibid.

33 Ibid 28.

34 Ibid 92.

35 Karl Marx, *Critique of the Gotha Programme*, in Karl Marx and Friedrich Engels: Collected Works vol 24, (Lawrence & Wishart 2010) 87.

36 See also Karl Marx, *Grundrisse: Foundations of the Critique of Political Economy* (Penguin Books 1993) 81–115.

37 Marx (n 35).

38 Paul O'Connell, 'Capitalism, Inequality, and Human Rights' (*Law and Political Economy*, 4 June 2018) <https://lpeblog.org/2018/06/04/capitalism-inequality-and-human-rights>.

reducing the distance between rich and poor, but if and how rights as a practice and discourse can contribute to the elimination of inequality by contesting the very roots of this inequality, namely, the structural reproduction of capitalist relations of production.

The limitations of Moyn's critique of inequality account for his misplaced urge for a return to a mid-twentieth-century model of welfare state, which would tackle the inequality which was exacerbated in the age of neoliberalism. To be fair, Moyn does not idealise the welfare state, since he recognises the role played in its establishment by the emergence of the Soviet Union as the victor of the Second World War, as well as the role of the welfare state as an alternative to revolution.[39] However, he fails at the same time to identify the precarious nature of welfare in capitalism and its contingent nature upon the balance of forces between social classes. This in turn explains his misplaced trust in the collectivist spirit of reconciliation at the heart of the welfare state, as well as his equivocal treatment of more confrontational rights, such as the right to strike or the right to collective bargaining,[40] in favour of rights such as the right to work or the right to the product of one's labour.

I argue that Moyn ultimately fails to identify the fundamental contradiction at the core of rights in capitalism because he fails to conceive of the struggle over rights as a form of class struggle. To illustrate this point let us briefly discuss the contradictory nature of right to work in capitalism. Moyn identifies the right to work (i.e. the obligation on government and society to provide gainful employment if none was available) as 'the first right in importance' after the French Revolution.[41] Moreover, the inclusion of the right to work in the Universal Declaration of Human Rights[42] is one of the reasons for Moyn's positive treatment of this document and the canon of ideals it consecrated.[43]

Nevertheless, despite the social and ontological significance of the category of labour[44] in human history, it is extremely doubtful that a right to work in a capitalist society has chances of ever being satisfied or enforced; in fact, historical evidence would prove otherwise. The reproduction of conditions for labour exploitation as well as unemployment are structural characteristics of capitalism and necessary for its reproduction, not the result of a lack of legal enforcement of the right to work; which is why in the first Constitution of a country that initiated the process of socialist construction (i.e. the 1918 Russian Soviet Federated Republic Constitution), 'work' appears not as a right but as a 'universal obligation' for the purpose of 'eliminating the

39 Moyn (n 9) 44.

40 Ibid 32.

41 Ibid 26.

42 Article 23.1 of the Universal Declaration of Human Rights states: 'Everyone has the right to work, to free choice of employment, to just and favourable conditions of work and to protection against unemployment.'

43 Moyn (n 9) 3, 6–7, 57–61.

44 Labour as a dialectical category plays a fundamental part in G F W Hegel's analysis of the process towards achieving self-consciousness. The process following the struggle to the death and the establishment of the positions of Master and Slave between two individual subjects is distinguished for its materiality and its capacity of revealing the transient nature of these social relations. It is through the process of labour that the state of Lord and Bondsman is negated; see G W F Hegel, *Phenomenology of Spirit* (Oxford University Press 1977) 118–119. In Karl Marx's dialectics, the importance of labour is furthered as it is linked with the concept of the totality of social relations. Man becomes separated from the animal world when he begins to work using implements of labour which he himself created. Thus, for the dialectical analysis, the real universal basis of everything that is human in man is production of instruments of production; see E V Ilyenkov (Aakar Books 2008) 74. Production of labour implements is the objective basis for all other human traits without exception, as the simplest, elementary form of man's human being; see Ilyenkov, ibid 75. It is for this reason that Georg Lukacs in his 'Ontology of Social Being' analyses the concept of labour as the model for social practice; see Georg Lukacs, *Ontology of Social Being: Labour* (Merlin Press 1980).

parasitic strata of society and organising the economic life of the country'.[45] This is evidence of the different approach to the place of labour in society in different social systems and its reflection in law.

In capitalist societies there is a contradiction inherent in the right to work, and different class interests from different social class standpoints can be reflected in it and facilitated by it. This right can be useful for the struggle of the toiling classes in capitalism, but it can also be filled with a content that facilitates exploitation against the interests of workers (as well as pensioners, injured or – even terminally – ill individuals) so as to contribute to the reproduction of the conditions of exploitation. One such example is to be found in the introduction of the principle of flexibility in labour relations, as part of the process of their deregulation.[46] Flexibility is nominally aimed at countering unemployment and, by extension, satisfying the right to work for a wider part of the population. However, the goal of reducing unemployment in reality stands for the true goal of reducing labour costs, through the intensified exploitation of a wider labour force. The reduction of unemployment, in this context, in actuality means the enhancing of the numbers of the reserve army capable of work, so as to lower the cost of labour. Part-time, temporary relations (as well as the introduction of educational schemes for the unemployed) favour the inclusion of previously excluded elements in the workforce, so that the abundance of supply and the increase of workers' exploitation reduce the labour costs. A right to 'flexible' work, thus, translates into measures which promote labour exploitation, through part-time and temporary contracts, performance-related wages, elimination of collective bargaining and facilitation of dismissals.

A similar contradiction appears in a recent resolution passed by the European Parliament on 'pathways for the reintegration of workers recovering from injury and illness into quality employment'.[47] This resolution could be seen as giving effect to the right to work of those ill and injured. This, of course, in itself is a valid goal. People have a right not to be discriminated against because of their illness or old age; but this has to be accompanied by a right to good working conditions, a right to training and a right to support. Nevertheless, in this instance, the right to work may be also employed in the favour of capital. Those who exercise their right to work cannot simultaneously exercise their right to pension or social support. The reintegration of injured or – even terminally – ill workers is part of the bourgeois governments' and supranational institutions' response to the problem of increase in the average life expectancy; a problem which is viewed as an obstacle to the profitability of capital, as is evident in the explanatory statement of the resolution.[48] Despite its nominal defence of the right to work, this resolution can be seen as part of the more general process of the dismantling of social security systems throughout EU Member States, so as to enable their subsequent privatisation.

45 Article 1, Chapter 2(f) of the 1918 Russian Soviet Federated Republic Constitution.

46 Flexibility is effected through policies which focus on 'removing obstacles which make it more difficult or costly to employ part-time workers or workers on a fixed-duration contract, and gearing careers more closely to the individual, or facilitating forms of progressive retirement'; on 'reducing working hours in a period of recession'; on 'gearing levels of pay to company performance and productivity'. See EU Commission, *White Paper on Growth, Competitiveness, and Employment* (COM (93) 700, 1993).

47 European Parliament resolution of 11 September 2018 on pathways for the reintegration of workers recovering from injury and illness into quality employment (2017/2277(INI)).

48 Ibid: 'Longer life expectancy combined with the increasing average age of retirement leaves Europe and Member States with significant challenges not only for our health systems but also for our labour markets . . . People with disabilities make dependable employees with comparable productivity, lower accident rates and higher job retention compared to company's general workforce. They represent an untapped source of skills and talent.'

D'Souza's critique of rights

The above examples point towards a conclusion regarding 'human' rights that forms the first thesis put forward here: *'human' rights can be recuperated by social agents other than those they seem to be designed for*. This thesis seems to be understood and well reflected in D'Souza's analysis of the 'human' right to happiness. Much like its originator (i.e. the inalienable right included in the US Declaration of Independence), the 'human' right to happiness is an abstract right which acquires its individualist connotations and bourgeois content once embedded in the framework of the dominant social relation of capital in the last few centuries. If understood as the right to leisure for workers and the toiling classes, or their right to vacations and a salary that can satisfy cultural and general development, the right to happiness can be seen as a carrier of important social demands. However, according to D'Souza, in the epoch of imperialism, the 'human' right to happiness is 'a calculated strategy for expanding the tourism and related industries by relying on legal treaties and health and welfare legislation in European Union member states'.[49]

D'Souza's argument becomes even more powerful when it manifests how this first thesis can be even further 'stretched' into the second: *'human' rights can be used crudely to legitimise imperialist interventions*. This point is one of the central tenets of D'Souza's critique of rights and is illustrated with several case studies. First and foremost, reference is made to the central role of 'democracy promotion' in US foreign policy.[50] The promotion of 'democracy' by the US government has a long history dating back to Woodrow Wilson's presidency (1913–1921): it marked the beginning of the Cold War by legitimising the intervention in the Greek Civil War[51] and became an overt foreign policy from the mid-1970s.[52] D'Souza describes how democracy promotion has over the last few decades been 'freed' from state monopoly and contracted out to a variety of social actors, like private foundations, NGOs, humanitarian organisations, think tanks, etc.[53]

Under this prism, D'Souza discusses the phenomenon of international election monitoring and the 'right to free and fair elections' in the epoch of imperialism. She begins by identifying how the right to self-determination and the principle of non-interference in the internal affairs of a state in international law meant that the political, diplomatic and military intervention by the US-led imperialist bloc in Third World countries had little legitimacy.[54] An antithesis, thus, emerged between self-determination, on the one hand, and the promotion of democracy and 'human' rights, on the other. The point of departure for the supporters of the latter was the view that the right to self-determination had 'become a facade for states to violate "human" rights'. When states act in their narrow self-interest and compromise ethics and universal norms, the UN must intervene to promote 'human' rights within states even if it means weakening the principle of self-determination.[55] Hence the role of the 'right to free and fair elections' in international law so as to legitimise the interference in the internal affairs of states. D'Souza refers to the 2005 report, entitled 'In Larger Freedom' by the UN General

49 D'Souza (n 8) 12.

50 Ibid 27–35.

51 On account of which the Truman doctrine was stated in front of Congress on 12 March 1947 by President Harry S Truman, who sought for financial and military 'assistance' for the Greek 'democratic' government which was 'threatened by the terrorist activities of several thousand armed men, led by Communists'.

52 D'Souza (n 8) 29.

53 Ibid 31.

54 Ibid 101.

55 Ibid 80.

Assembly, which sought to modify the sovereignty principle by imposing a duty on the 'international community' to intervene to monitor democracy.[56]

D'Souza furthers this argument with a discussion on the antithetical approaches to the right to self-determination between US President Wilson and V I Lenin. On the one hand, Wilson promoted the idea of *legal* rights in the name of freedom for colonised nations; formalised later in the 1933 Montevideo Convention on Rights and Duties of States, between the US and Latin American states, and extended under the Atlantic Charter, between the US and the UK in 1941.[57] For Wilson, the agents of freedom in the colonies were the financiers and investors who would bring modernisation to the colonies through their investments. Thus, according to D'Souza, the purpose of self-determination as a *legal* right was 'to establish equal access to colonial markets between capitalist states'.[58] On the contrary, according to Lenin, self-determination could never be a *legal right*, but was instead a *political claim* which colonies had to fight for through political actions. According to D'Souza, this theoretical understanding moved many national liberation struggles away from demands for legal rights, channelled their energies into organising for political change and, by bringing the socialist and national liberation struggles closer, extended moral and material support to anti-colonial movements.[59]

This brings us to the third thesis which I consider as flowing directly from D'Souza's analysis: *'human' rights can be used to divert social struggles away from political goals and canalise them into legal proceedings*. It can be argued that 'human' rights contribute to the reproduction of capitalist social relations by contributing to the process of depoliticisation and individualisation of social problems; that is by turning a political issue into a technical legal issue and a matter of collective struggle into a matter of individual struggle.[60] This is why, according to D'Souza, there is an important lesson to be learned from the anticolonial struggles of socialists and freedom fighters, who demanded food and not the right to food, national independence and not the right to independence:[61] to demand not the legal protection of a right, but the satisfaction of the social need itself.

The dangers of framing wide social and political demands in the language of rights are well understood by O'Connell in his critique of the displacement thesis. This thesis maintains that reliance on the language of 'human' rights by movements for radical social change is problematic because it tends to crowd out (or displace) other, potentially emancipatory, languages, and as a consequence distract attention from broader, structural causes of injustice and oppression. O'Connell disagrees with this nihilistic approach towards 'human' rights. He argues that the mobilisation of rights language can make an important contribution to movements for radical social change, without displacing or

56 Ibid 96.

57 Ibid 192.

58 Ibid.

59 Ibid.

60 It has been argued that the process of depoliticisation of social and economic issues is essential to the process of reproduction of capitalism. See E M Wood, *Democracy Against Capitalism: Renewing Historical Materialism* (Verso 2016); Louis Althusser, *On the Reproduction of Capitalism* (Verso 2014). Marx's insight in Capital was that he understood that the main conditions of surplus appropriation in capitalism (i.e. the separation of the individual producer from the conditions of labour and the absolute private property of the means of production) rest on a specific and historically conditioned political configuration and force-relation between classes. Depoliticisation (i.e. the separation of the economic from the political) means that these conditions are presented by bourgeois political economists as natural conditions and, as a result, not subject to political debate or contestation. It is, therefore, important to consider the ways in which the 'human' rights discourse contributes to this.

61 D'Souza (n 8) 210.

precluding the mobilisation of other emancipatory languages, and the challenging of deeper, structural causes of injustice. In support of his argument, he uses the examples of the Focus E15 campaign and the Right2Water movement in Ireland, which 'engaged the language of human rights alongside other frames of reference, and connected the specific rights struggle its protagonists were engaged in with broader causes of injustice'.[62]

It is possible here to discern the different nuances between the two authors' arguments. Both recognise the limits of the 'human' rights discourse, their individualising and depoliticising effect, as well as their aptness for recuperation by social forces for reasons other than those they seem to be designed for. However, D'Souza argues that, in the epoch of imperialism, rights 'lose any limited progressive potential they may have had in the nineteenth century',[63] whereas O'Connell argues that the human rights discourse can operate alongside other critiques against commodification, privatisation and austerity. Based on their analyses the following sections in this paper will examine ways in which the rights discourse may contribute to the struggle for strategic political and wider social goals.

Rights and social struggle

This section will continue with the next three theses which distinguish between rights and 'human' rights for the reasons stated in the first section, namely the rejection of the individualist connotations of the bourgeois conception of humanity. Instead, I argue that rights can only properly be understood on the basis of the contradictory relations of capitalism. This is reflected in the fourth thesis advanced here: *the struggle over rights is one form of class struggle*. This was the central point in Marx's analysis of the class struggle over and the legislative intervention on the regulation of the working day. For Marx, the capitalist's right as a purchaser to make the working day as long as possible is countered by the worker's right as a seller to reduce the working day to a particular normal length:

> There is here therefore an antinomy, of right against right, both equally bearing the seal of the law of exchange. Between equal rights, force decides. Hence, in the history of capitalist production, the establishment of a norm for the working day presents itself as a struggle over the limits of that day, a struggle between collective capital; namely the class of capitalists and collective labour (i.e. the working class).[64]

I argue that this point applies in general to the establishment of norms that regulate labour relations. The institution and abolition, protection, restriction or circumvention of rights is contingent on the balance of social forces and the development of class struggle. For instance, labour rights were not recognised in the early days of the bourgeois state. It was not until after the Second World War that most Western bourgeois states recognised labour rights. This recognition can be (to a large extent) attributed to the class struggle of the working class and the popular strata. Similarly, the attack on social rights, labour legislation and the welfare state, in these same countries over the last decades, can be (to a large extent) attributed to the class struggle of the capitalist class to get rid of these rights so as to restore its profitability through intensified exploitation of labour.

Let us illustrate this point with reference to the right to strike. The right to strike itself is now a legal (and in certain jurisdictions a constitutional) right whose recognition was the result of bloody struggles on behalf of the toiling classes. Strike actions were

62 O'Connell, 'Human Rights: Contesting the Displacement Thesis' (n 10) 27–32.

63 D'Souza (n 8) 206.

64 Marx (n 1) 344.

criminalised in the nineteenth century. They were born as a means for the workers of the new bourgeois industrialised societies to struggle for the improvement of their working conditions; intolerable conditions which resulted in accidents with hundreds of dead and injured in mines and factories. In Britain a strike was called as early as 1842 as part of the Chartist movement for fair wages and working conditions. In Greece, strike actions were criminalised until 1920, when recognised by Act N 211/1920.[65] In 1975 the right to strike was recognised and enshrined in the Greek Constitution. Article 23, paragraph 2, provides that: 'Strike constitutes a right to be exercised by lawfully established trade unions in order to protect and promote the financial and the general labour interests of working people.' The right is protected unconditionally. The only limits to the protection of the right relate to the subject of its exercise and its aims.

Nevertheless, the right is not wholeheartedly accepted by bourgeois society and its juridico-political apparatus. The unconditional protection of the right in the Constitution was followed by article 19 of Act N 1264/1982 which established further conditions for the legitimate exercise of the right (24-hour notice to the employer, provision for security staff, declaration of strike by the competent organ). A literal and teleological interpretation of the constitutional provision would consider that these conditions lie beyond the scope of the constitutional provision and are, therefore, inapplicable. On the contrary, not only are these legislative conditions applied by the courts but based on these the vast majority of strike actions are declared illegal. As a matter of fact, from 2009 to 2014 – during the first hard years of the last economic crisis – 300 judicial decisions on the legality of strike actions had the following outcome: 264 were declared illegal or abusive; 10 of the applications were dismissed for formal reasons; only 26 strikes were declared legal.[66] So, 88 per cent were declared illegal and only 8 per cent legal.

Another way through which the constitutional protection of the right to strike is circumvented in Greece is the governmental prerogative known as 'civil mobilisation' (πολιτική επιστράτευση) which can be used to terminate a legally called and organised strike action. The decree in question (17/1974) contains a very loose definition of 'emergency' which is apt for targeting industrial action as it is meant to cover events that can bring about damage to life or property or 'cause disruption to the economic and social life of the State'. Furthermore, the decree does not specify any time limits for its application, which allows the government to declare a sector under a civil mobilisation regime for months or even years. For instance, in June 2013 the subway employees were mobilised and the ban to strike action was lifted one year later (July 2014). In general, between 2010 and July 2014, the government exercised its powers under the decree to bring a strike to an end six times. During the first six months of 2013 alone, the decree was used three times.[67]

In the aftermath of the crisis, further limits to the exercise of the right to strike have been legislated following the EU principle of 'best practice'. Article 211 of Act N 4512/2018 limits the exercise of the right to strike by setting a requirement for a 50 per cent turnout of the union's registered members. This measure followed the 'best practice' already applied in Britain. According to section 2 of the Trade Union Act 2016, a 50 per cent turnout of those entitled to vote is required for the decision for industrial action to be valid. In parallel, in important public services, a further requirement of 40

65 Christophoros Sevastidis, *The Right to Strike and the Judicial Review of its Exercise* (in Greek) (Sakkoulas Publications 2015) 7.

66 Ibid 9.

67 Ntina Tzouvala, 'Continuity and Rupture in Restraining the Right to Strike' in Honor Brabazon (ed), *Neoliberal Legality: Understanding the Role of Law in the Neoliberal Project* (Taylor & Francis 2016).

per cent support is added, rendering it even more difficult for industrial action to be taken in sectors of health, education, the fire department, transportation, etc. It is evident from the above that the conditions of exercise of the right to strike – a right so vital for the needs and the negotiating power of workers – are worse than the previous regime which required a simple majority of those present in the ballot.

Of course, these measures are not an exception to the EU institutional apparatus; they are rather a manifestation of the general approach of the EU towards socio-economic rights. In the landmark cases *Laval* (C-341/05) and *Viking* (C-438/05) the Court of Justice of the EU held that the right to take industrial action is fettered in so far as it restricts freedom of movement and freedom of establishment such that, where industrial action restricts freedom of movement or establishment, it will only be lawful if it is both justified and proportionate. This approach is indicative of the capitalist orientation of the EU and reflects the interests and views of colossal capitalist institutions, such as JP Morgan. In fact, a report conducted by JP Morgan Chase in 2013 on the process of adjustment of the Euro-area economies to the crisis, southern European Constitutions and their constitutional protection of labour rights are seen as aberrations to the EU social acquis and as obstacles to growth and competitiveness.[68]

It is, thus, safe to argue that attacking the right to strike is a necessary aspect of bourgeois class struggle, especially during periods of crisis and intensified contradictions, because it removes a 'weapon' – perhaps the most powerful one – from the worker's 'armoury'. Additionally, restricting the right to strike may result in the hindering of the processes of class organisation[69] and class consciousness.[70] It is for this reason that the EU, bourgeois governments and the international capitalist system carry out a sustained attack on this right so as to reproduce or facilitate favourable conditions for capitalist exploitation.

---

68 David Mackie, Malcolm Barr, et al, 'The Euro Area Adjustment: About Halfway There' (*European Economic Research/JP Morgan*, 28 May 2013) <www.europe-solidarity.eu/documents/ES1_euro-area-adjustment.pdf>.

69 The example of the October 2018 ballot for a strike action by the University and College Union (UCU) is illustrative here. The October 2018 ballot manifested the full effect of the 50 per cent turnout requirement set by the Trade Union Act 2016. The previous ballot (January 2018) on a strike action to defend pensions had resulted in a 58.3 per cent overall turnout (24,707 votes cast in a total number of 42,145 members in the institutions concerned). This mandate gave effect to a strike action that lasted for several months. One result of this action and the subsequent mobilisation of workers in higher education was the increase in the numbers of union members by 16,000. However, this positive development with regards to class organisation was counteracted by the provision of the 2016 Act. To give an indicative example: at Birkbeck College, University of London, the amount of votes cast in the two ballots was almost identical (223 in the January ballot versus 222 in the October ballot); however, in the first ballot this amounted to a 53.7 per cent turnout, whereas in the second ballot only to a 42.05 per cent turnout, because union membership in this institution had risen from 415 to 528. All relevant information can be found at the UCU website <www.ucu.org.uk/media/9730/HE-pay-and-equality-industrial-action-ballot---full-results-Oct-18/pdf/ucu_he-ballot-report_oct18.pdf> and <https://www.ucu.org.uk/media/9091/USS-ballot-results---ranked-summary-table/pdf/uss_ballotresults_summaryranked_jan18.pdf>.

70 The educative effect of strikes in workers' consciousness has been highlighted in radical thought. See, for instance, Rosa Luxemburg, *The Mass Strike* (Bookmarks 1995); and V I Lenin, 'On Strikes' in *Lenin: Collected Works* vol 4 (Progress Publishers 1977). It has been argued that a strike action teaches workers about their real power and makes them realise their role in the production process. Even more importantly perhaps, a strike action teaches the workers to move beyond their immediate interest. As Lenin puts it, 'every strike means many privations for the working people, terrible privations that can be compared only to the calamities of war – hungry families, loss of wages, often arrests, banishment from the towns where they have their homes and their employment'; see Lenin, ibid 315. The sufferings during the strikes are definitely not in the immediate interest of the individual worker. But, the goal of improving working conditions and increasing wages is mediated by a process which goes against their immediate interests. The workers thus learn to differentiate between the immediacy of self-interest and the mediated consciousness of class interest.

This brings me in turn to the fifth thesis: *the protection, restriction or elimination of rights may serve either the immediate or the strategic interests of social classes*. I argue that the relationship between rights and social struggle is mediated through social and class interests. It is in the pursuit of their class interests that workers go on strike and, consequently, fight for the right to go on strike. Similarly, it is in the pursuit of their class interest that the bourgeois class seeks the restriction of this right.

Further to this, I argue that there is a distinction between what we may call an immediate interest and what may be called a mediated or a strategic interest. We find an example of this in Marx's analysis of the struggle for the length of the working day. The lengthening of the working day and the increase of labour exploitation is in the interest of the capitalist, as it directly increases capital's profitability. However, the restriction of the working day so as not to exhaust the worker and reproduce their labour power is also in the interest of capital. For Marx, the English Factory Acts were the negative expression of capital's 'appetite for surplus labour' and 'blind desire for profit'.[71] Such laws are in the strategic interest of capital. The longevity of labour exploitation by capital is mediated through the restriction of the working day and the limiting of capital's immediate profitability.

The same point can be raised *mutatis mutandis* regarding labour law and the welfare state in general. It can be argued that the welfare state was in the strategic interest of capital. It served the reproduction of capitalism by facilitating the expanded reproduction of social capital following the Second World War. Furthermore, the welfare state and the recognition of socio-economic rights served the strategic interest of reproduction of capital by absorbing the social movements that had threatened the European ruling classes for centuries. Therefore, the argument can be sustained that the welfare state, labour law and socio-economic rights, despite on the face of it contradicting the interests of the capitalist and intervening in the field of class struggle in support of the worker, *strategically* may serve the interests of capital.

Conversely, it is in the workers' *immediate* interest to ameliorate the conditions of exploitation; to improve their working conditions so as to eliminate accidents; to increase their wages; and to reduce the working day. This is why the toiling classes carry out their class struggle. But, if the *amelioration* of the conditions of exploitation is in the *immediate* interests of workers, it is the *elimination* of these conditions of exploitation that constitutes their *strategic* interest. To the extent that labour law, socio-economic rights and the welfare state can advance some of the workers' demands in capitalism and improve their working and living conditions, they serve their immediate interests. But, to the extent that these institutions contribute to the reproduction of capitalism by not contesting the fundamental conditions of exploitation (i.e. the private ownership of the means of production), they do not serve the strategic interest of the working class and popular strata.

I can now turn to the sixth and final thesis: *the struggle of the working class and the popular strata for rights may only serve their strategic interest (i.e. the elimination of the conditions of exploitation) if it is ultimately directed towards the elimination of the capitalist social relations*. The above analysis highlighted the importance of socio-economic rights, and the right to strike in particular, as well as the vital role of the struggle for these rights. However, it also highlighted their precariousness in capitalism. The status of social rights in capitalism is precarious and the onslaught against those rights over the last four decades is proof of this. Does this mean then that we have to eliminate the discourse of rights from social struggles? I claim that

71 See Marx (n 1) 348: 'These laws curb capital's drive towards a limitless draining away of labour-power by forcibly limiting the working day on the authority of the state, but a state ruled by capitalist and landlord.'

the recognition of the limitation of the struggle for rights in capitalism should not amount to an outright rejection of this struggle and of the role that the rights discourse can play in the working-class struggle. I argue instead for the necessity of re-evaluating the role of rights in social struggle. What is necessary is to struggle for rights while at the same time recognising the limitations of any demands and forms of struggle articulated in the rights discourse.

I would argue that any critique and struggle for social justice articulated within the rights discourse has to be expanded and assume a materialist content. As mentioned above, the amelioration of the conditions of capitalist exploitation is limited compared to the elimination of those same conditions. Similarly, addressing the issue of distributive relations in capitalism is limited compared to addressing the issue of productive relations in capitalism, namely, the root of the imbalance of forces between different social classes as they participate in the production process. Recognition of the limitations of the rights discourse may result in recognition of the need to transform a struggle framed in this discourse into a struggle against the totality of capitalist social relations; that is transform an individual struggle for a legal right into a collective struggle for collective demands. Furthermore, it may necessitate the move from a critique of distributive relations in capitalism to a critique of capitalist productive relations, since the former are themselves conditioned by the latter.[72]

Last, but not least, expanding the rights discourse would certainly necessitate a move beyond the juridification of social struggle. The above analysis of the right to strike showed that the institutionalisation of a social goal and the constitutional enshrinement of a right is not the end of the process or its ultimate goal. It is, thus, essential to move beyond positive law and avoid limiting a social struggle in the strict confines of juridification as a goal. The juridification of the struggle may reflect a temporary victory of the workers' struggle and serve their immediate interests, but it has to be remembered that the legal status of rights – social and economic rights of the toiling class in particular – in capitalism is precarious and vulnerable to all sorts of mechanisms aiming at the restriction or elimination of these rights.

So, the goal is the unity of the struggle *for* rights with the struggle *against* capitalism – the unity of economic and political class struggle. To unite the economic and the political class struggle is essential in order to overcome in practice the separation of the economic from the political sphere in capitalism and, thus, raise consciousness of the deeply political nature of everyday problems in capitalist society. As we saw above, class struggle carried out in the form of a strike action is significant for the process of the development of class consciousness. For instance, with 88 per cent of strike actions declared illegal by the Greek courts, it is safe to argue that a strike action may have a rather educative effect in the workers' consciousness regarding the role of the state in bourgeois society. Additionally, the fact that a strike action necessarily teaches the workers to move beyond their immediate interest is another example of how social struggle may contribute to the process of development of class consciousness, as well as to a radicalisation of the demands that carry the struggle forward.

Therefore, expanding a critique of inequality framed in the language of rights into a comprehensive critique of social relations in capitalism can be a pathway for uniting the struggle for every-day matters, such as working and living conditions, with the struggle for a different social arrangement altogether that prioritises social needs instead of profit. Ultimately, this move beyond juridification of social struggle is an essential aspect of the

72 See Marx (n 35).

process of transforming the struggle for the distribution of wealth and the satisfaction of social needs to a struggle for radical change of the mode of production which involves a thorough and comprehensive critique of the extant power, property and productive relations.

Conclusion

To conclude, this article examined several aspects of the relationship between rights and social struggle. Reviewing recent literature on this topic it outlined six theses that characterise this relationship. Starting with Moyn's intellectual and ideological history of 'human' rights, it highlighted its limitations. It was argued that Moyn's analysis of inequality in capitalism does not concern itself with the structural roots of inequality. It focuses solely on distributional relations in capitalism while neglecting the issue of exploitation in the sphere of production. It therefore fails to conceive the struggle for rights as a form of class struggle and misevaluates the role of rights in addressing social inequality.

On the contrary, D'Souza's critique of rights presents a solid base for assessing the role of rights in the epoch of imperialism because it examines rights in the context of the struggle between different social forces. Her analysis provides a strong argument that rights can be used for other purposes and by other social agents than those they seem to be designed for; in fact, they have been used to justify imperialist interventions. D'Souza assesses negatively the role of rights – and especially 'human' rights – in the epoch of imperialism and concludes that social movements ultimately need to demand not the legal protection of a right, but the satisfaction of the social need which this right addresses. O'Connell, on the other hand, while recognising the limitations of the rights discourse, argues that 'human' rights language can be used alongside other critiques of social injustice.

Acknowledging the difficult questions raised by these thinkers and their fruitful contributions, this article was structured around one central question: to what extent can the rights discourse promote a comprehensive critique of capitalist social relations? I argue that the rights discourse can be an essential element in the working-class struggle only to the extent that it is expanded, meaning only as long as it teaches the toiling classes to move from a struggle for rights to a struggle against the totality of capitalist social relations; to move from quantitative change (concessions in capitalism) to qualitative change (supersession of capitalism) – to put it in more abstract terms.

The rights discourse can be the spearhead of a comprehensive critique of social injustice once it is expanded into a critique of capitalist social relations as a whole; that is once the idea of rights itself assumes a distinctively materialist content instead of an idealist one. For this, a different conception of justice is necessary: a materialist conception of justice. A conception, the essence of which is captured in the slogan sung by the Greek workers in the strike actions in the years of the crisis: 'The workers' justice is the law' ('νόμος είναι το δίκιο του εργάτη'). I argue that this slogan contains the seeds for a materialist conception of justice and can be viewed as a maxim (*jus proletarii suprema lex esto*) that contains a normative statement: a normative demand for the elimination – and not just the amelioration – of the conditions for labour exploitation.

NILQ 70(4): 425–444

Causing controversy: interpreting the requirements of causation in criminal law and tort law

GEMMA TURTON

University of Sydney Law School

AND

SALLY KYD

Leicester Law School *

Abstract

The occurrence of a fatal road traffic collision may raise a number of legal issues and result in litigation both in the civil and criminal courts. The role of the different branches of law is distinct, with the aims of the litigation being quite different, but both require causation to be proved. Such cases are examined in this article as a vehicle for discussing how the principles of causation play out in each branch of law. It will be seen that the particular aims of the law dictate how doctrines of causation are applied, with particular problems caused by the legislature's creation of strict liability offences. To resolve these problems, we propose that the criminal law borrow from negligence in adopting a test akin to the 'harm within the risk' test, adapting it to the role of the criminal law by formulating a 'harm within the wrong' requirement for causation.

Keywords: causation; tort; negligence; criminal law; harm; wrong; driving; responsibility

1 Introduction

One of the most common events in modern life likely to lead to involvement by a member of the public with both the criminal and civil justice systems of England and Wales is a road traffic collision (RTC). In 2018 there were 160,378 people killed or injured on Britain's roads as the result of a collision.[1] The injured party, or bereaved relative, might seek compensation for their injuries or loss through the civil courts under the tort of negligence, while concurrently any negligent driver might be charged with an applicable criminal offence and prosecuted in either the magistrates' court or the Crown Court. The two branches of law will be engaged for two different purposes, to be explored throughout

* Gemma Turton is Senior Lecturer at the University of Sydney Law School; Sally Kyd is Professor of Law at Leicester Law School. We would like to thank the reviewers for their constructive comments on earlier drafts, as well as a number of other individuals who commented on previous drafts, including Findlay Stark, Matt Dyson, Sandy Steel and Sarah Green. An early version of this paper was presented at the Society of Legal Scholars 2014 annual conference, and we wish to thank Leicester Law School for financial and study leave support in the development of this article.

1 Department for Transport, *Reported Road Casualties in Great Britain: Main Results 2018* (Department for Transport 2019) <www.gov.uk/government/statistics/reported-road-casualties-great-britain-main-results-2018>.

this article, but at the same time they are, as recently noted by Dyson, 'tightly bound together'.[2]

Let us take a not atypical scenario involving a fatal RTC to frame the question we seek to address. According to Clarke et al, right of way violations are the cause of a significant proportion of collisions.[3] Such collisions typically involve a number of factors converging at once to produce tragic consequences. These factors might include a failure to abide by the rules of the road by more than one road user: typically a car driver who fails to look properly in order to ensure that the road across which they intend to turn is clear of other traffic, and a motorcyclist travelling towards the car in excess of the speed limit.[4] In such a case, whether the Crown is able to secure a conviction for causing death by careless driving[5] will depend upon two main factors: that D's driving fell below the standard of a competent and careful driver, thus amounting to careless driving,[6] and that D's careless driving *caused* V's death. Whether the deceased's estate is able to sue for damages in negligence will depend on similar questions: did D breach a duty of care owed to V? The standard of care applied to determine whether D breached such a duty is similar to that under the criminal law of driving without due care and attention (careless driving): the standard of the reasonably competent driver.[7] Secondly, did the breach of duty *cause* the death? The different contributions to the collision can be represented in a reduction of damages due to contributory negligence.[8] If V was speeding, this is likely to provide a reduction of around 50 per cent,[9] although in the worst cases D's contribution might be reduced to as little as 20 per cent where the claimant's speed was particularly excessive.[10] In a criminal prosecution for a causing death offence, however, the issue is one of 'all or nothing' liability. D is either liable for causing death, or they are not.[11]

What if V is 100 per cent to blame for the RTC? Careless driving on D's part could not be proved in such a case, but what of an offence of causing death by unlawful driving under s 3ZB Road Traffic Act 1988, which provides that a person is guilty of an offence if he causes the death of another person by driving a motor vehicle on a road when he is unlicensed or uninsured?[12] The offence is one of strict liability, in that the defendant need not be aware that they were unlicensed, uninsured, or disqualified; neither need they have any *mens rea* in relation to causing death or even creating the risk of any harm. Much, then,

2 M Dyson, 'Disentangling and Organising Tort and Crime' in M Dyson (ed), *Unravelling Tort and Crime* (Cambridge University Press 2014) 20. For an alternative approach to comparing causation in the two branches see A P Simester, 'Causation in (Criminal) Law' (2017) 133 Law Quarterly Review 416.

3 In a sample of fatal RTCs examined to establish their cause, 16 per cent involved right of way violations, leading Clarke et al to identify this as one of two problematic areas for road safety, the other being loss of control collisions, making up 44 per cent of fatal RTCs: D Clarke, P Ward, C Bartle and W Truman, 'Killer Crashes: Fatal Road Traffic Accidents in the UK' (2010) 42 Accident Analysis and Prevention 764, 768.

4 Such examples can be seen in an empirical study of prosecutions brought as a result of fatal road traffic collisions: S Cunningham, 'Has Law Reform Policy Been Driven in the Right Direction? How the New Causing Death by Driving Offences are Operating in Practice' (2013) 9 Criminal Law Review 712.

5 This is the most likely charge to be brought, under s 2B Road Traffic Act 1988. See Cunningham (n 4).

6 Road Traffic Act 1988, s 3ZA.

7 *Nettleship v Weston* [1971] 2 QB 691 (CA).

8 Law Reform (Contributory Negligence) Act 1945, s 1(1).

9 See, for example, *Wheeler v Chief Constable of Gloucestershire Constabulary and Others* [2013] EWCA Civ 1791.

10 See, for example, *Ringe v Eden Springs (UK) Ltd* [2012] EWHC 14 (QB); *Train v Secretary of State for Defence* [2014] EWHC 1928 (QB).

11 Although it is recognised that the contribution of the victim's own negligence can be taken into account as a mitigating factor in sentencing.

12 Until 2015 it also included causing death by driving whilst disqualified, but a separate offence was created by s 29 of the Criminal Justice and Courts Act 2015, which inserted s 3ZC into the Road Traffic Act 1988.

depends on the interpretation of the requirements of causation in such a case, an issue that was considered by the Court of Appeal in *R v Williams*[13] and *R v Hughes*,[14] before being clarified by the Supreme Court.[15] As will be seen, *Hughes* may have reduced the injustice created by the offence being one of strict liability by incorporating a fault requirement into causation, but it remains the case that a defendant may be convicted under s 3ZB in circumstances where she would not have been liable to compensate the victim in the tort of negligence. This paper considers the question of what it means 'to cause death by driving' from the perspectives of both criminal law and negligence.[16]

There are a number of aspects of criminal law that result in criminal liability where no civil liability is incurred, such as the meaning of appropriation in the law of theft,[17] or the unavailability of defences such as consent or *ex turpi causa* which could lead D to avoid liability in tort.[18] Indeed, attempts liability exists in criminal law – but not in negligence, where damage is required – since criminal law is concerned with censuring wrongs, including the wrong in second-order harms such as threats to security even, in relation to some inchoate crimes and endangerment offences, where specific 'victims' may be unidentifiable. As Lord Scott stated in *Ashley v Chief Constable of Sussex Police*, '[a] plea for consistency between the criminal law and the civil law lacks cogency for the ends to be served by the two systems are very different'.[19] Dyson summarises these ends:

> . . . criminal law responds to moral, natural or public wrongs with a penalty and should be characterised by clear, certain and formal rules; tort law responds to many of the same moral and natural wrongs, but does so for private parties and tends to focus on putting the wrong right, rather than punishing.[20]

This paper is concerned with causation and the question of how the purposes of the two systems of law are reflected in their respective causal doctrines. Causation is intimately tied up with notions of responsibility, indeed in *Environment Agency v Empress Car*, Lord Hoffmann stated that 'one cannot give a common sense answer to a question of causation for the purpose of attributing responsibility under some rule without knowing the purpose and scope of the rule'.[21] Our focus, then, is on the place of causation in these different systems of responsibility. The offence and facts in *Hughes* raise thorny questions about causation in the context of strict liability and where the victim was also at fault. We address these in section three of the paper where we identify the need for a more nuanced approach to legal causation where liability is strict, and in the final section we advocate developing a 'harm within the wrong' test to keep criminal liability within appropriate bounds. First though, it is helpful to articulate some of the more basic divergences in approach to causation and to situate them within the purposes of the two fields of law. In part this is because, while these divergences are commonly known, they

13 *R v Williams* [2010] EWCA Crim 2552, [2011] 1 WLR 588.

14 *R v Hughes* [2011] EWCA Crim 1508, [2011] 4 All ER 761.

15 *R v Hughes* [2013] UKSC 56, [2013] 1 WLR 2461.

16 Unless otherwise stated, any reference to 'negligence' refers to the tort of negligence under civil law.

17 *R v Hinks* [2001] 2 AC 241. For a comparison of the crime of theft and the civil wrong of conversion, see S Green, 'Theft and Conversion – Tangibly Different?' (2012) 128 Law Quarterly Review 564. For other comparisons between the civil and criminal law, see Dyson (ed), *Unravelling Tort and Crime* (n 2).

18 G R Sullivan, 'Wrongs and Responsibility for Wrongs in Crime and Tort' in M Dyson, *Unravelling Tort and Crime* (ed) (n 2) 85.

19 *Ashley v Chief Constable of Sussex Police* [2008] UKHL 25, [2008] AC 962, [17].

20 M Dyson, 'Tortious Apples and Criminal Oranges' in M Dyson (ed), *Comparing Tort and Crime: Learning from Across and Within Legal Systems* (Cambridge University Press 2015) 421.

21 *Environment Agency v Empress Car Co (Abertillery) Ltd* [1999] 2 AC 22 (HL), 31.

are rarely articulated side by side, and in part it helps contextualise the analysis of the specific problems of strict liability that later arise.

2 Divergences in approach to causation in criminal law and negligence law

Depending on one's perspective, the law on causation in negligence is more fully developed than in criminal law, or as Sullivan describes it 'causation in the criminal law is a less cluttered field than is the case in tort law'.[22] The purpose of this section is to identify the differences in basic approach to causation in the two fields and to posit some possible explanations. Our focus is on the approaches to factual causation and *novus actus interveniens*, but for clarity's sake we begin by outlining the general approach to causation in each area of law. A complete account of causation in each field is beyond the scope of this work, but a broad outline helps locate key points of divergence.

A thread common to both branches of law is the distinction drawn between issues of factual and legal causation. Glanville Williams explains the nature of these inquiries:

> When one has settled the question of but-for causation, the further test to be applied to the but-for cause in order to qualify it for legal recognition is not a test of causation but a moral reaction. The question is whether the result can fairly be said to be imputable to the defendant . . . If the term cause must be used, it can best be distinguished in this meaning as the 'imputable' or 'responsible' or 'blameable' cause, to indicate the value-judgment involved.[23]

We will return to this concept of moral responsibility later; for now, it suffices to note that the distinction between factual causation and legal causation, in both branches of law, broadly corresponds to the division between factual issues and evaluative questions about responsibility. Factual causation is established in negligence using the but-for test,[24] or the *Wardlaw* test of material contribution to harm,[25] and legal causation comprises the *Wagon Mound* test of reasonable foreseeability of the type of harm as the test of remoteness,[26] subject to the thin skull principle that the extent of harm need not be foreseeable,[27] and the *novus actus interveniens* doctrine. In criminal law, factual causation is a narrower test determined using the but-for test alone;[28] then at the legal causation stage courts ask whether there was a *novus actus interveniens*, meaning that this but-for cause was no longer substantial and operating[29] at the time of the relevant harm.[30]

Causation in negligence is more developed in two key senses. First, the tests of factual causation (and development of exceptional tests) have been the subject of a number of judicial decisions and academic discussion in recent years.[31] Second, as Sullivan highlights,

22 Sullivan (n 18) 102.

23 G Williams, *Textbook of Criminal Law* (2nd edn, Sweet & Maxwell 1983) 381. For the need of the legal system to make this distinction in relation to a road traffic fatality specifically, see A Merry and A McCall Smith, *Errors, Medicine and the Law* (Cambridge University Press 2001) 132.

24 *Barnett v Chelsea and Kensington Hospital Management Committee* [1968] 2 WLR 422 (QB).

25 Exceptionally the *Fairchild* test of material contribution to the risk of harm may apply, but this test is confined to very limited cases: *Fairchild v Glenhaven Funeral Services Ltd* [2002] UKHL 22.

26 *The Wagon Mound (No 1)* [1961] AC 388.

27 *Smith v Leech Brain & Co Ltd* [1962] 2 QB 405.

28 The usual authority cited for this rule is *White* [1910] 2 KB 124.

29 D Ormerod and K Laird, *Smith and Hogan's Criminal Law* (14th edn, Oxford University Press 2015) 95. A number of cases can be cited as authority for this test, including *R v Smith* [1959] 2 QB 35.

30 See discussion below.

31 Key examples are the *Fairchild* exception: *Fairchild* (n 25); developments of the *Wardlaw* test such as *Bailey v MOD* [2008] EWCA Civ 883 and *Williams v Bermuda Hospitals Board* [2016] UKPC 4; and *Chester v Afshar* [2004] UKHL 41.

'[b]roadly, under criminal law, questions of whether an outcome was foreseeable and not too remote . . . are not entertained at the stage of causation/*actus reus*'.[32] The evaluative work is done instead by *mens rea*, a point which acquires greater significance when we later consider strict liability offences where, of course, there is not a *mens rea* requirement. We now consider possible explanations for these divergences.

2.1 Factual causation

In relation to factual causation one might expect the law in both areas to align; since evaluative issues are primarily the realm of legal causation, the different objectives of liability in the two fields would naturally acquire greater significance there. In negligence, courts have faced increasingly complex factual issues, particularly in the industrial disease and medical contexts. These are arguably inescapable given the central role causation plays in connecting the claimant to the defendant in a system of interpersonal responsibility and have led to the factual/legal causation divide becoming more pronounced. It might also be suggested that negligence law has developed more detailed and tightly focused tests for the more practical reason, identified by Hart and Honoré, that causation serves wider functions in negligence than in criminal law. In criminal law generally 'a causal connection between some action of the accused and the specified harm must be shown in order to establish the *existence* of liability', while 'in tort causal questions are usually relevant both to the existence of liability and to its extent'.[33] Since the extent of the harm caused by the negligence determines the quantum of damages for which D is liable, it is natural that negligence law more finely delineates the relationship of causation.

This impacts on the approach to causation where the defendant's act or omission merely accelerates an outcome. In *R v Morby*, the father failed to seek medical treatment for his son who died of smallpox, and his conviction was quashed because it was not proven that the omission 'caused death or accelerated it'.[34] In Simester's view, acceleration of the outcome would not be a sufficient basis to conclude that the father caused the child's death if treatment would ultimately have failed.[35] In similar circumstances, however, negligence liability may be imposed (where a duty of care is owed) but, as noted earlier, causal considerations affect not only the existence of liability but also its extent; the claimant's already shortened life-expectancy can be taken into account at the quantification stage.[36] Following *Williams v Bermuda Hospitals*,[37] a negligent delay in treatment may materially contribute to the outcome and therefore be a cause when it permits the claimant's condition to worsen, even if on the balance of probabilities the ultimate outcome is unchanged. This was not an uncontroversial decision,[38] but on orthodox principles, if the delay in treatment itself causes damage in the form of worsening the condition or causing a longer period of pain and suffering prior to diagnosis then the defendant may be held liable for that if not the final outcome.[39] The flexibility that exists in negligence to conclude that the defendant's omission caused part of the total loss or accelerated its occurrence allows for a wider range of contributions

32 Sullivan, (n 18) 102.

33 H L A Hart and Tony Honoré, *Causation in the Law* (2nd edn, Oxford University Press 1985) 8485.

34 *R v Morby* (1882) 8 QBD 571, 575 (Stephens J) cited in Simester (n 2) 437.

35 Simester (n 2) 437.

36 See e.g. S*mith v Leech Brain & Co Ltd* [1962] 2 QB 405.

37 *Williams v Bermuda Hospitals Board* [2016] UKPC 4.

38 See e.g. S Green, 'Q: When is a Material Contribution not a Material Contribution? A: When It Has not Been Proven to have Made any Difference to the Claimant's Damage' (2016) 32 Professional Negligence 169.

39 *Hotson v East Berkshire HA* [1987] AC 750 (HL).

to be treated as causal than may be the case in criminal law. On the other hand, the 'all or nothing' nature of liability in criminal law makes the meaning of causation of great importance, so one could equally expect the criminal law to have developed clear and consistent principles as to its application.

A further possible reason for the divergence is identified by Shute: he argues that the reason why the (criminal) law tends to regard causation in terms of broad generalisations based on common sense principles rather than to mimic the more obscure approach to causation often taken by the philosopher or scientist (or perhaps the tort lawyer) is that in criminal law *jurors* need to understand causal principles.[40] This is, of course, true of cases triable on indictment, and, for offences triable summarily or either way, the fact-finders who must understand the law are often non-legally qualified lay magistrates, as opposed to the legally experienced judge making decisions as to causation in negligence cases. A related point is that the burden of proof is different in criminal cases compared to civil cases. In a criminal case, the fact-finders can only convict if they are 'sure' of the accused's guilt (including that D caused the relevant harm in the case of a result crime), whereas in a civil case the claimant's case only need be proved 'on the balance of probabilities'. Beyond these differences, we now turn to consider the causal issues that arise, and how they are addressed, when the victim was also at fault.

2.2 Causation and the victim's own carelessness

The victim's own carelessness introduces causal complexities in both criminal law and negligence, and, while it goes to the existence of liability in both fields through the doctrine of *novus actus interveniens*, in negligence it may affect the extent of liability through the defence of contributory negligence. This, we suggest, reflects the different aims of punishment and allocation of responsibility and affects how issues of *novus actus* are resolved in each field.

To contextualise the discussion of the victim's own carelessness, where a *novus actus* is the act of a third party, in both negligence and criminal law it is worth noting that liability is not binary. If a subsequent act of a third party, such as a medical practitioner, does not break the causal chain, that third party does not necessarily escape liability. She may commit a criminal offence such as gross negligence manslaughter[41] and, in negligence, may be held jointly and severally liable along with the defendant for the additional damage caused. Conversely, the simple fact that the third party commits a criminal offence or is negligent does not entail that she breaks the chain of causation in criminal law and negligence respectively. Indeed, given that negligent medical treatment is a foreseeable and not uncommon event, it will generally not break the chain of causation in either field. If anything, criminal law appears to take a stricter approach to subsequent medical negligence, requiring treatment to be so bad as to be so independent and so potent in causing death as to be a supervening act,[42] while in negligence it seems that gross medical negligence would break the chain of causation resulting in D avoiding liability for the additional damage.[43] On one hand, it seems strange that negligence law will more readily accept that medical treatment breaks the chain of causation than criminal law since, where joint and several liability applies, this recognises the responsibility of both defendants. However, criminal law censures not only the causation of harm but the

40 S Shute, 'Causation: Foreseeability v Natural Consequences' (1992) 55 Modern Law Review 584, 584.

41 Applying the test in *Adomako* [1995] 1 AC 171.

42 *R v Cheshire* (1991) 93 Cr App R 251.

43 *Webb v Barclays Bank* [2001] EWCA Civ 1141, [2002] PIQR P8.

intention to cause it; given the moral luck that causation involves,[44] it is understandable that it should be confined to a small role with more significance placed on the need to punish the blameworthy assailant in addition to the blameworthy medic.

The contrasting rationales behind negligence and criminal law become more significant where the potential *novus actus* is an act of the victim. While the criminal liability of the defendant remains all or nothing, in negligence the responsibility for the damage can be shared between the claimant and defendant through the defence of contributory negligence. The conclusion in a criminal trial is thus a rather stark one to be determined by the fact-finders: D is either 'guilty' or 'not guilty' of the offence charged.[45] Where guilt is established, Simester argues that 'the scalarity of causation should directly affect sentencing, in as much as differences in one's degree of causal responsibility would imply differences in one's degree of culpability'.[46] However, it is still true to say that criminal liability is 'all or nothing' in the sense that if the victim's conduct constitutes a *novus actus*, liability will not ensue. A *novus actus* of the claimant would similarly result in the defendant escaping liability in negligence, but the availability of the contributory negligence defence means that the role of the claimant can more readily be taken into account and lead to apportionment of damages where the claimant acted negligently, albeit not in a manner that amounted to a *novus actus*. Thus, in the decision in *Scott v Gavigan*, the Court of Appeal considered it preferable that the claimant's negligence in a road traffic collision should constitute contributory negligence rather than a *novus actus* in all but 'pretty exceptional circumstances'.[47] Clarke LJ explained that because it is not uncommon for someone to run out into the road, this is a risk that drivers should foresee. Since they therefore '[owe] a duty to take care not to injure even the foolish, I find it difficult to see why [they] should be absolved of all liability and the claimant denied any relief save in extreme circumstances'.[48] The primary question for criminal law in such circumstances is whether D's act of driving is sufficiently blameworthy to attract liability for the end result. The difference is that V's own negligence could not influence the outcome of the case, unless it were in fact extreme enough to absolve D of liability.[49] Steel argues that 'if . . . the rules on intervening agency are the legal precisifications of the broad question "Whose doing

44 Simester states that what he labels 'mechanical' causation is the 'primary vehicle for allocating moral luck': Simester (n 2) 423.

45 Although it is true that one should not overemphasise the role of the fact-finders in determining the conclusion, in the sense that the outcome of a criminal case depends upon a number of factors, not least the selection of offences to be charged by the Crown and the decision of D to plead guilty or not guilty to such charges or lesser included offences.

46 Simester (n 2) 424.

47 *Scott v Gavigan* [2016] EWCA Civ 544 [34].

48 Ibid.

49 In *Scott v Gavigan* (ibid) the V suffered a 'significant injury' to his leg. In that situation, a criminal court would not be called upon to decide whether D had 'caused' the injury unless D's driving was considered dangerous, and so he might be liable for an offence under s 1A Road Traffic Act 1988, or if he were disqualified from driving, in which case he might be liable for the offence under s 3ZD Road Traffic Act 1988. Otherwise, D might be liable, not for the end result, but for an endangerment offence such as careless driving, if it was thought that his failure to avoid V amounted to driving which fell below the standard of a competent and careful driver, in which case V's injuries are irrelevant to the charge. Parliament's approach has been rather piecemeal in introducing constructive result crimes from underlying conduct crimes, and it is this which has created the headache for courts in determining issues of causation in such cases. Following its most recent consultation on driving offences, the government intends to bring forward proposals to create a new offence of causing serious injury by careless driving: Ministry of Justice, *Response to the Consultation on Driving Offences and Penalties Relating to Causing Death or Serious Injury* (Cm 9518, 2017) <www.gov.uk/government/uploads/system/uploads/attachment_data/file/651879/consultation-response-on-driving-offences.pdf>.

was this?'', then we would not expect the rules to vary between tort and crime'.[50] Tort law, however, is much more readily able to handle the answer that it was the doing of *both* the claimant and defendant since the aim is allocation of responsibility rather than blame or censure. This may explain, for example, why in *Pagett* police officers who shot towards D, killing the girlfriend he was using as a human shield, were seen to be acting involuntarily out of self-preservation and in the line of duty and did not break the chain of causation between D's act and V's death, despite being liable to V's mother in civil law.[51]

Additionally, negligence law tends not to ask such an open-ended question as 'Whose doing was this?'. Instead, the causation inquiry is tightly framed by prior conclusions about damage, duty and breach, so that we ask whether the defendant's failure to take reasonable care, in breach of a duty owed to the claimant, caused the damage that the claimant suffered. In criminal law, causation is constitutive of the *actus reus* of an offence so bears the burden of tasks that, in negligence, would be addressed by duty and breach. Many evaluative questions about responsibility and blame are traditionally addressed within *mens rea*. In the absence of a *mens rea* requirement, those evaluative questions remain to be addressed, but, we will argue, it is inappropriate to attempt to subsume them within the causation inquiry.

The case of *Hughes* raises the difficult question of the weight to be given to the victim's own carelessness when D's liability is strict, so is not premised on carelessness on D's part. The defendant in *Hughes* had been driving his camper van uninsured when he was confronted with the victim's oncoming car, which veered onto D's side of the carriageway and collided with the camper van. It was found that the deceased had been under the influence of drugs and overtired, and it was accepted that D's driving was faultless and that he could have done nothing to avoid the collision.[52] The relevant offence in this case was s 3ZB Road Traffic Act 1998 according to which D is guilty of an offence if he causes the death of another person by driving a motor vehicle on a road and, at the time when he is driving, he is uninsured. Since the victim did not voluntarily and deliberately kill himself this was not a case of *novus actus interveniens* but of concurrent causes, and the question was whether the driving of D was a cause of V's death. The Supreme Court went on to hold that in order to find that D so caused the death of V there must be 'at least some act or omission in the control of the car, which involves some element of fault . . . which contributes in some more than minimal way to the death', and on the facts this fault was missing.[53] While it is understandable that the court would seek to introduce an element of fault into a homicide offence, it will be argued that the way in which this was done does not provide a coherent basis for criminal liability that can be of general application.

In determining liability, one issue that troubled the court was that V 'was not an innocent victim and could never have recovered any compensation if he had survived injured'.[54] Although the facts of *Hughes* are typical of those that could potentially give rise

50 S Steel, 'Causation in Tort Law and Criminal Law: Unity or Divergence?' in M Dyson (ed), *Unravelling Tort and Crime* (n 2) 272.

51 *R v Pagett* (1983) 76 Cr App R 279. See A Norrie, *Crime, Reason and History: A Critical Introduction to Criminal Law* (3rd edn, Cambridge University Press 2014) 185. The High Court awarded £8000 to V's mother several years later (3 December 1990); presumably what counts as 'reasonable' will vary between self-defence in criminal law and the standard of care in negligence.

52 Other than to not have been driving at the time and on the road to be involved in a collision.

53 *R v Hughes* [2013] UKSC 56, [36].

54 *R v Taylor* [2016] UKSC 5, [9]. The certified question also made specific reference to circumstances where the manner of the defendant's driving was faultless and 'the deceased was (in terms of civil law) 100% responsible for causing the fatal accident or collision', at [35].

to negligence liability, we should not allow this to cloud our thinking by encouraging us to analyse the issues arising from the case in negligence terms. While it might be thought that of the two branches of law it is criminal law, not tort, which ought to be the more restrictive, the reason there is no negligence liability on the facts of *Hughes* is simply that the defendant had not breached a duty of care towards the victim. This outcome does not turn on causation because in the absence of a breach of duty there is no causation question to be asked. While causation is constitutive of *actus reus* in criminal law, in negligence liability the causation question is tightly framed by the damage and breach of duty, so causation questions simply do not arise in the absence of a breach, for example, failure to drive with reasonable care.

There is therefore a limit to how useful it was, given the current position, for the Supreme Court to highlight the different outcome in negligence when addressing the causal requirement in criminal law. If the court's concern was simply that there may be criminal liability in circumstances that would not attract negligence liability, then, as we have seen, there are sometimes good reasons for differences in the law. If the court's concern instead was that the s 3ZB offence is primarily concerned with the absence of insurance and the financial harm caused to claimants who are subsequently uncompensated, then it may be concerning that criminal liability can arise where there would be no tortious liability, but this is a concern which should be directed at the drafting of the legislation. A more coherent approach would have been for Parliament to add the absence of insurance as an aggravating factor to the causing death by careless driving offence, along the lines of the offence under s 3A of the Road Traffic Act 1988, which carries a maximum sentence of 14 years' imprisonment where death caused by careless driving is aggravated by D driving under the influence of alcohol or drugs.[55] Moreover, the Supreme Court's solution does not avoid this risk, since criminal liability will now arise under s 3ZB where there is something properly to be criticised in the manner of D's driving, which still does not amount to careless driving so may also not amount to negligence.

What the offence under s 3ZB highlights in particular is the idea that it is the balance that needs to be struck between the role played by moral luck, on the one hand, and blameworthiness, on the other, that is perhaps key to determining the demands placed on causation in criminal law. Many of the leading cases on causation in criminal law are cases of constructive or 'unlawful act' manslaughter,[56] where the harm actually caused might be far out of proportion to the blameworthiness of D. Lord Goff's recognition in *Pagett* that the fact that causation is present will not necessarily lead to a conviction for murder or manslaughter, since other elements of the offence such as *mens rea* must be proved, is tempered by his admission that 'in the majority of cases he is likely to be guilty at least of manslaughter'.[57] Given the ease with which unlawful act manslaughter can be proved, in that D need have no awareness of the risk of death or harm she poses to her victim, unlawful act manslaughter has come under fire from subjectivist criminal law commentators who see the breach of the principle of correspondence (requiring that the *mens rea* of an offence relate to the proscribed harm resulting from the *actus reus*) and the

55 It is by no means suggested that a maximum of 14 years would be proportionate for such an aggravated offence, however. It can also be argued that such a separate offence would not be required, given that lack of insurance operates as an aggravating factor in sentencing: Sentencing Guidelines Council, *Causing Death by Driving: Definitive Guideline* (Sentencing Guidelines Council 2008) <www.sentencingcouncil.org.uk/wp-content/uploads/Causing-death-by-driving-definitive-guideline-Web.pdf>.

56 For more in-depth discussion of this point, see S Kyd, 'Causing Death' in M Bohlander and A Reed (eds), *Homicide in Criminal Law: A Research Companion* (Routledge 2018).

57 *R v Pagett* (n 51) 289.

emphasis on the role of luck as problematic.[58] Objectivists, on the other hand, see little problem with finding D liable for causing death where she has made her own moral luck.[59] That does, however, result in an increase in the importance of the clarity of the law on causation, which is currently lacking.[60]

Moral luck encompasses luck as to a range of matters: the outcomes we cause, the circumstances we find ourselves in and our own disposition. Nagel advanced the view that moral judgements cannot be isolated from the effects of luck; moral judgements are sensitive to factors that we have no control over.[61] The concept of causation is, clearly, important to identifying outcomes as 'ours', distinguishing those outcomes that we can be said to have caused despite the role of luck from those that we cannot be said to have caused. But, as Ashworth identifies, we need to separate from causal responsibility 'certain questions of moral responsibility: is she morally to blame for the accident? Does he have a moral duty to compensate the owner of the vase?'[62] In negligence, a degree of stigma will inevitably flow from a finding of liability, and this will vary depending on the damage caused, which does entail a role for outcome luck, yet the remedy remains compensation and the tort remains negligence. In contrast, Ashworth argues, 'the criminal law is chiefly concerned with desert, that is, with whether or not the person deserves to be labelled as a criminal and, if so, what level of offence is fairly applicable . . . culpability is a crucial issue'.[63] Where there is luck as to the outcome caused, this potentially affects not only the existence of liability in criminal law, but also the offence for which D is convicted, with the labelling function affecting the degree of stigma attached and the severity of the punishment imposed. A minor assault can become manslaughter if the victim dies, even where the risk of death was not foreseeable. Given that criminal law invites a wider range of moral judgements to be made against the defendant than negligence law, we might expect outcome luck to be confined to a smaller role through more restrictively applied causal concepts. Instead, it seems to be negligence law that allows a smaller role for outcome luck by requiring that the type of harm suffered would have been reasonably foreseeable.

Horder seeks to defend unlawful act manslaughter against concerns about the increased role it gives to moral luck, arguing that 'the existence of an intention to do wrong may make it legitimate to hold someone criminally responsible for any adverse consequences of which there was a risk in committing the intended wrong, whether it could be said that the risk would reasonably foreseeably turn into a reality or not'.[64] He distinguishes between two defendants who discharge guns close to the victim who suffers a shock and consequently dies of a heart attack. The defendant whose gun went off while he was unlawfully cleaning it has probably not committed manslaughter but, Horder argues, the defendant who deliberately fired his gun near to the unsuspecting victim is in a morally different position because he has made his own luck by directing his efforts

58 See, for example, A Ashworth, 'Change of Normative Position: Determining the Contours of Culpability in Criminal Law' (2008) 11 New Criminal Law Review 232; B Mitchell, 'More Thoughts about Unlawful and Dangerous Act Manslaughter and the One-punch Killer' [2009] Criminal Law Review 502.

59 See, for example, J Horder, 'A Critique of the Correspondence Principle in Criminal Law' [1995] Criminal Law Review 759.

60 See, for example, *Carey* [2006] EWCA Crim 17; commentary of *Carey* by D Ormerod: [2006] Criminal Law Review 842.

61 T Nagel, *Mortal Questions* (Cambridge University Press 2013) 24–38.

62 A Ashworth, 'Taking the Consequences' in S Shute, J Gardner and J Horder (eds), *Action and Value in Criminal Law* (Clarendon Press 2013) 112.

63 Ibid 117.

64 Horder (n 59) 764.

towards harming the victim, so is morally responsible for the death.

Negligence liability clearly does not require intention to harm, but the conduct giving rise to negligence liability is 'directed' towards the claimant in as much as the defendant must have owed a duty of care to the claimant and must have exposed the claimant to an unreasonable risk of harm. The conduct underlying the offence in s 3ZB is the strict liability offence of driving whilst uninsured. It is not 'directed' towards harming the victim. Since the defendant need have no knowledge that she is uninsured, she has not altered her moral position vis-à-vis the victim. It is the outcome, involving outcome luck, that is primarily being invoked to justify a moral judgement and criminal liability for the death. This increases the demands that are placed on causation, and in the final section we now consider how best to address these.

3 Causation and strict liability: a 'harm within the wrong' test for strict liability result crimes

The decision in *Hughes* has not been confined to the s 3ZB offence; in *R v Taylor*, the Supreme Court held that it extends to the offence under s 12A(2)(b) of the Theft Act 1968[65] under which a person is guilty of aggravated vehicle-taking if, owing to the driving of the relevant vehicle, an accident occurred by which injury was caused to any person.[66] The basic offence of taking a vehicle is not one of strict liability, in that it requires knowledge of the absence of authority. However, the aggravated offence is constructive in nature and does not require explicit proof of a legally recognised species of *mens rea* as to an aspect (the resulting harm) of the offence,[67] despite being a serious crime carrying a maximum sentence of two years' imprisonment, or 14 years if it causes death.[68] Again there was no fault in the manner of D's driving of the stolen vehicle in *Taylor*,[69] and the wording of the statute is more neutral in using the language of 'owing to the driving' rather than 'causes … by driving', yet the Supreme Court held that there must be something properly to be criticised in the manner of D's driving. Throughout both judgments there are references to this being the normal common law approach to causation,[70] or a common-sense meaning of causation,[71] but, as we will discuss, the court elides causal responsibility with moral responsibility or blame. This section addresses two preliminary issues: the rationale for strict liability in tort and crime and the relationship between causal, moral, and legal responsibility. Following this discussion, we question how causation should be established for such liability under the two branches of law, ultimately arguing that adoption of a 'harm within the wrong' test in strict liability result crimes would be preferable to altering the meaning of causation more generally in order to bridge the gap between causal and moral responsibility.

65 Inserted by s 1 Aggravated Vehicle Taking Act 1992.

66 *R v Taylor* [2016] UKSC 5, [2016] 1 WLR 500.

67 Duff's definition of formal strict liability: R A Duff, *Answering for Crime* (Oxford University Press 2007) 233.

68 In *R v Sherwood & Button* (1995) 16 Cr App R (S) 513 it was found that the different sentences provided under s 12A(4) Theft Act mean different offences are created, including one of 'causing death'.

69 Although D was drink-driving, the court focused narrowly on the manner of his driving and found that there was no fault.

70 *R v Hughes* [2013] UKSC 56 at [16], [20], [27]; *R v Taylor* (n 54) [18].

71 *R v Taylor* (n 54) [23], [25].

3.1 Divergence in justifications of strict liability

The offences in both *Hughes* and *Taylor* were crimes of strict liability. Despite the conventional view that criminal liability is constructed through the commission of an *actus reus* committed with *mens rea*, there are an increasing number of regulatory offences that depart from this paradigmatic construction in relation to one, or other, or both requirements. In recent years, regulatory offences have come to dominate the criminal law,[72] and these are often created as crimes of strict liability. Driving is one form of activity that is regulated through the criminal law in this way. Many driving offences are conduct crimes regulating the act of driving where no harm need result, meaning that causation is not an issue that need be addressed, and no knowledge of the wrong committed need be proved. This can be contrasted with offences of pollution where, again, knowledge of the wrong is not needed, but it must be proved that D's act or omission caused the pollution and is thus a result rather than conduct crime. These strict liability result crimes are difficult to fit into a rational exposition of the law, and the causal issues they raise are often addressed under an additional subheading by textbook writers in addition to the categories of breaks in the causal chain discussed in the previous section.[73] The purported advantages of crimes of strict liability are certainty, efficiency and effectiveness.[74] In *Taylor* the court observed that strict liability in criminal law is generally 'founded on collective convenience rather than moral imperatives' and 'although fault in the actual commission of the offence may be unnecessary, there are nonetheless positive steps that the prospect of criminal liability may cause people to take in order to prevent the offence from occurring'.[75]

In comparison, strict liability is not widespread in tort law,[76] and 'the strictness of the liability varies considerably along a spectrum from near absolute liability to little more than a reversed burden of proof'.[77] The rationale for strict liability is less clear-cut in tort, but Peel and Goudkamp suggest 'it is perhaps possible to discern behind some of them a very hazy idea of unusual or increased risk'.[78]

The regulatory criminal offence of driving without insurance is concerned not with the risk of physical damage but with the *financial* harm of road-users being unable to obtain compensation for damage or injury resulting from road traffic 'accidents'.[79]

72 Chalmers and Leverick found that 98 per cent of the 1395 offences created in 1997–1998 and 89 per cent of the 1760 offences created in 2010–2011 were regulatory in nature, applying to those acting in some form of special capacity: J Chalmers and F Leverick, 'Tracking the Creation of Criminal Offences' [2013] Criminal Law Review 543, 558.

73 'Two special cases of intervention: pollution and driving' is a sub-heading used in A P Simester, J R Spencer, F Stark, G R Sullivan and G J Virgo, *Simester and Sullivan's Criminal Law: Theory and Doctrine* (7th edn, Hart 2019) 108. '*Empress Car*' is a subheading used in D Ormerod and K Laird, *Smith and Hogan's Criminal Law* (14th edn, Oxford University Press 2015) 99.

74 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, 693. For a more nuanced defence of strict liability, see A P Simester, 'Is Strict Liability Always Wrong?' in A P Simester (ed), *Appraising Strict Liability* (Oxford University Press 2005).

75 *R v Taylor* (n 54) [26].

76 In the sense of 'outcome-based' strict liability, which Cane distinguishes from strict liability in the sense of simple interference with an interest of the claimant, and relationship-based strict liability in the form of vicarious liability: P Cane, 'Retribution, Proportionality, and Moral Luck in Tort Law' in P Cane and J Stapleton (eds), T*he Law of Obligations: Essays in Celebration of John Fleming* (Clarendon Press 1998) 150.

77 W E Peel and J Goudkamp, *Winfield and Jolowicz on Tort* (19th edn, Sweet & Maxwell 2014) 28.

78 Ibid.

79 A point made by Lords Hughes and Toulson at para [9] in *Hughes* (n 15).

Stevens argues that regulatory offences are 'not dependent upon there being either an interpersonal wrong or want of moral virtue' so it is 'unsurprising that the offences can be committed blamelessly'.[80] The s 3ZB offence is premised on a regulatory offence which is not justified by the risk of physical harm but the risk of such harm going uncompensated, yet it is ultimately a result crime so should depend upon an interpersonal wrong or want of moral virtue. While convenience and efficiency may be sufficient to justify regulatory offences 'especially where the offences involved are common and individually involve no serious harm or even any harm at all',[81] this rationale cannot extend to a homicide offence which by definition involves serious harm.

3.2 Distinguishing types of responsibility: causal, moral, and legal

What does it mean to say that an individual is responsible for another's death? The debate here turns on the boundaries between causal responsibility, moral responsibility and legal liability,[82] and the discussion of causation in *Hughes* does not adequately distinguish these concepts. Admittedly, the distinction between causal and moral responsibility is muddied in the law by the fact that the causation inquiry comprises both factual causation and legal causation, so legal causation introduces issues of moral responsibility into the realm of 'causation'.[83]

Hart famously set out a range of forms of responsibility, distinguishing role responsibility, causal responsibility, legal and moral liability responsibility, and capacity responsibility. Notably, causal responsibility 'need not carry even an implication that [a person] was deserving of censure or praise; it may be purely a statement concerned with the contribution made by one human being to an outcome of importance, and be entirely neutral as to its moral or other merits'.[84] Beever similarly distinguishes between causal responsibility and moral responsibility, using the following example:

> Imagine that I am crossing a bridge on my way home from work. Though I have no reason to suspect it, the bridge has a structural weakness. The outcome of this weakness is that my being on the bridge causes it to collapse. Though I survive, the collapse causes the deaths of 10 people.[85]

He goes on to explain that he can accept causal responsibility for the deaths, but that does not mean he is personally morally responsible for them.[86] He suggests that there is no reason to think that it is inappropriate for the person crossing the bridge to think that this is '*something that happened to him* rather than something for which he was personally responsible'.[87] This echoes Ashworth's concern about the role of outcome luck in the constructive crime of unlawful act manslaughter, that 'the fault and the result are simply

80 R Stevens, 'Private Rights and Public Wrongs' in M Dyson (ed), *Unravelling Tort and Crime* (n 2) 143.

81 Simester et al (n 73) 211.

82 For a discussion of these boundaries in the context of criminal responsibility and liability, see Duff (n 67).

83 See J Stapleton, 'Unpacking Causation' in P Cane and J Gardner (eds) *Relating to Responsibility: Essays in Honour of Tony Honoré on his 80th birthday* (Oxford University Press 2001); J Stapleton, 'Choosing What We Mean by "Causation" in the Law' (2008) 73 Missouri Law Review 433, 446.

84 H L A Hart, *Punishment and Responsibility: Essays in the Philosophy of Law* (2nd edn, Oxford University Press 2008) 215. See also Simester (n 2) 416 and 422–433; A Beever, 'Corrective Justice and Personal Responsibility in Tort Law' (2008) 28 Oxford Journal of Legal Studies 475, 489.

85 Beever (n 84) 489.

86 Ibid 490.

87 Ibid, original emphasis.

too far apart for a manslaughter label to communicate anything other than the misfortune which befell both the victim and [D]'.[88]

In criminal law, in relation to the s 3ZB offence, we can say that causal responsibility rests on D in that D's driving of the car was a *sine qua non* connection with V's death, and a *sine qua non* connection *is* a causal connection. The subsequent narrowing that takes place to identify the *responsible* cause involves a value judgement dependent on context and the purpose of the inquiry. Fault is not a requirement for causation, but fault is often a requirement for moral responsibility and legal liability. Fault is not part of the character of causation, it is part of the character of the conduct to which the law will apply a test of causation.

Negligence lawyers and criminal lawyers may well disagree over this proposition, and one explanation is the different place occupied by causation in establishing liability. As previously discussed, in negligence the causation question is tightly framed by the damage and breach requirements, so the question is specifically whether D's breach of duty caused the damage suffered by the claimant, and it is simpler to separate the factual and moral issues. In criminal liability the causation inquiry is an element of establishing the *actus reus* so forms part of a more open-ended inquiry into what D did, where factual and moral issues are more tightly interwoven. However, in criminal law it is the requirement of *mens rea* which normally determines the presence of moral responsibility for serious result crimes. It is when *mens rea* is dispensed with that the problem of moral luck means that difficult questions are posed in relation to causation, since causation is required to do some of the work normally reserved for *mens rea* in establishing moral culpability for outcomes that were beyond D's control.[89]

3.3 Causation in strict liability in tort and crime

Assuming that it is accepted that a *sine qua non* connection is a causal connection, it is still necessary to consider how to frame the counterfactual question that establishes factual causation in strict liability and in negligence. This is more straightforward in negligence since it is focused on D's breach of duty, so the formulation of the breach shapes the counterfactual question. The actual world is compared with a hypothetical world where D did not breach her duty. It is accepted that this hypothetical is a world where D conforms to the relevant standard of care; she cannot escape liability by arguing that she would have breached her duty in some other way and thereby still have caused the harm.[90] Stapleton suggests that '[t]he defendant's behaviour is altered *just* enough to bring it into conformity with his duty as mandated by the Law'.[91] So, for example, where a defendant driver was speeding, the relevant hypothetical world is one where she was driving at a reasonable speed at the time of the accident. Hamer explains, '[i]t would be too much of a departure from the actual world to construct a possible world where the defendant was not driving at all'.[92] More precisely, Wright says that we remove the tortious aspect of D's conduct

88 Ashworth (n 62) 120.

89 See text to n 87 above.

90 *Bolitho v City of Hackney Health Authority* [1998] AC 232 (HL). Here, the doctor negligently failed to attend the patient who later suffered respiratory failure. She argued that, even if she had attended, she would not have intubated and that was the treatment that would have prevented respiratory failure. The House of Lords said the breach of duty test has to be applied to her hypothetical decision not to intubate; if that would also have been negligent then she is still a factual cause.

91 Stapleton, 'Choosing What We Mean by "Causation" in the Law' (n 83) 451, original emphasis.

92 D Hamer, '"Factual Causation" and "Scope of Liability": What's the Difference?' (2014) 77 Modern Law Review 155,164.

since this 'ensur[es] that the reasons for making a person subject to liability govern and limit the extent of her liability'.[93]

If there were a duty of care in negligence to have public liability insurance when driving, the absence of insurance is the tortious aspect that would change in the hypothetical world. Having insurance would not change the outcome; if D was driving with reasonable care then the harm would still have occurred even if she had been insured. As noted by Merkin and Steele, 'absence of insurance does not itself cause death'.[94] Thus, in *The Empire Jamaica*,[95] there was no causal link between the fact that an officer (the defendant) navigating a ship did not have the certificate required by law and the collision that occurred while he was navigating. Honoré explains that '[g]iven that the basis of liability was fault, it was rightly held that the lack of certificate was irrelevant, since the officer would have navigated no better had he possessed one'.[96]

In relation to *The Empire Jamaica*, Honoré goes on to argue, however:

> [H]ad there been strict liability for navigating the ship without a certificate, so that if the ship was navigated by a pilot without a certificate the navigation was at the defendants' risk, the navigation would have been held to have caused the collision. Since strict liability is liability not for wrongful conduct, but for engaging in risk-creating activity, there would have been no need in this case to show that the lack of certificate was causally relevant. In such a strict liability case, it would have been enough that, had the ship not been navigated by X (who did not in fact possess a certificate), no collision would have occurred.[97]

Effectively, he suggests that where liability is strict rather than fault-based, then the appropriate hypothetical *is* a world where the ship is not sailed at all, so for s 3ZB the relevant question is whether the fact of D's car being driven at all was a cause of V's death. In contrast to negligence law where the tortious aspect of D's conduct is substituted with reasonable conduct, Honoré explains that:

> Substitution of lawful conduct is not possible and is not required in cases of strict liability, since the defendant's conduct, though it creates a risk, is not unlawful unless it causes harm. In such a case the hypothetical inquiry thus must be whether the plaintiff would have suffered the injury had the defendant not engaged in the activity . . . that entails strict liability.[98]

While the defendant in *Hughes* was engaged in a risky activity, *viz* driving, the particular risk being addressed by the s 3ZB offence was the financial risk of victims being uncompensated when injured by uninsured drivers. It is less than clear that the relevant hypothetical is a world in which the uninsured defendant does not drive at all, and in the following section we suggest that a 'harm within the wrong' test would allow courts to openly engage with consideration of what the relevant hypothetical is in any particular strict liability result crime.

---

93 R Wright, 'The Grounds and Extent of Legal Responsibility' (2003) 40 San Diego Law Review 1425, 1495. See also R Wright, 'Causation in Tort Law' (1985) 73 California Law Review 1735, 1805–06.

94 R Merkin and J Steele, 'Policing Tort and Crime with the MIB: Remedies, Penalties and the Duty to Insure' in M Dyson (ed), *Unravelling Tort and Crime* (n 2) 36.

95 *NV Koninklijke Rotterdamsche Lloyd v Western Steamship Co Ltd (The Empire Jamaica)* [1955] 3 WLR 385.

96 T Honoré, 'Necessary and Sufficient Conditions in Tort Law' in T Honoré, *Responsibility and Fault* (Oxford University Press 1999) 101. Broadbent similarly argues that a hypothetical world where the ship is not sailed at all would involve a 'gratuitous departure from the facts': A Broadbent, 'Fact and Law in the Causal Inquiry' (2009) 15 Legal Theory 173, 190.

97 Ibid.

98 Ibid 105–106.

3.4 Future directions: a 'harm within the wrong' test

Ultimately, we agree with Simester that any discomfort in finding that *Hughes* caused death should be directed towards the legislator,[99] which should not have created a strict liability homicide offence lacking in *mens rea*. But, given that these offences exist and appropriately require something more than causal responsibility, it is essential to address the question of what that 'something more' is, and where it can be injected. In addition to the existing tests of legal causation, it might be argued that criminal law would benefit, as Stuckenberg suggests, from a kind of remoteness principle akin to the 'harm within the risk' approach which originated in tort law (again with Lord Hoffmann)[100] and asks whether the actual harm is of the kind which the violated rule was designed to prevent.[101] While Turton has criticised the risk principle in negligence,[102] the more relevant limitation of it in this context is that, as we have seen, strict liability in the criminal offences under consideration is not risk-based. A more appropriate test, we will suggest, would be a 'harm within the wrong' test. Focusing on the wrong further recognises the distinctive purpose for which the criminal law is acting, the purpose of expressing censure, and concretises Lord Hoffmann's assertion that one cannot determine issues of causation without first knowing the purpose and scope of the relevant rule.

Articulating the harm within the risk test, Lord Hoffmann gave the example of a mountain climber about to undertake a difficult climb who seeks medical advice about the fitness of his knee, and whose doctor negligently fails to diagnose a knee injury, meaning that the climber goes on to undertake a climb that they would not have undertaken had they known of the injury. Whilst climbing he suffers an injury that is a foreseeable consequence of mountaineering but unrelated to the fitness of his knee, such as being swept up in an avalanche. Although the doctor's negligence is a but-for cause of the climber's presence on the mountain and therefore of the subsequent injury, the scope of the doctor's duty of care does not extend to all mountaineering injuries, only those related to the fitness of the knee.[103] The Model Penal Code (MPC) of the USA seeks to apply a 'harm within the risk' test to criminal liability.[104] In relation to paradigmatic crime, the *mens rea* requirement (acting 'purposely or knowingly') will in most cases provide the basis of the causal relationship between the conduct and result. However, in relation to crimes of negligence and absolute (strict) liability, conditions are applied to the finding of causation:

Model Penal Code §2.03

(3) When recklessly or negligently causing a particular result is an element of an offense, the element is not established if the actual result is not within the risk of which the actor is aware or, in the case of negligence, of which he should be aware unless:

(a) the actual result differs from the probable result only in the respect that a different person or different property is injured or affected or that the probable injury or harm would have been more serious or more extensive than that caused; or

99 Simester (n 2) 441.

100 *South Australia Asset Management Corp v York Montague Ltd (SAAMCo)* [1997] AC 191 (HL).

101 C-F Stuckenberg, 'Causation' in M D Dubber and T Hornle, *The Oxford Handbook of Criminal Law* (Oxford University Press 2014) 484–485.

102 G Turton, 'Informed Consent to Medical Treatment Post-*Montgomery*: Causation and Coincidence' (2019) 27 Medical Law Review 108, 122–130.

103 *SAAMCo* (n 100) 214.

104 See also Kyd (n 56).

(b) the actual result involves the same kind of injury or harm as the probable result and is not too remote or accidental in its occurrence to have a [just] bearing on the actor's liability or on the gravity of his offense.

(4) When causing a particular result is a material element of an offense for which absolute liability is imposed by law, the element is not established unless the actual result is a probable consequence of the actor's conduct.

We can see here that, for offences requiring some form of *mens rea*, the test for causation is bound together with the state of mind of the defendant, according with psychologists' findings that ordinary people's judgement of causal selection may be affected by an independent judgment of blame.[105] The problem is that such judgements will be made in the same way in relation to offences of strict liability, which cannot be catered for by the law, in that the very point of the law is that the defendant need not be aware of any risk. As a result, the test for 'absolute' offences is fudged in the MPC, with a rather woolly requirement that the result be a 'probable consequence' of the actor's conduct. This appears rather unsatisfactory to adopt without amendment if we are to introduce something similar here.

We would, therefore, seek to adapt the rule and, rather than frame the requirement in terms of risk, we would, instead, employ the concept of the *wrong* that the offence seeks to address. The regulatory offence of driving without insurance is not a moral wrong in itself and, as such, amounts to *malum prohibitum* rather than *malum in se*,[106] which exist to regulate driving on a public road and to prevent financial harm. Duff gives the examples of speeding and gun laws in noting that large parts of the criminal law take the form of coercive regulation, punishing conduct unconnected to wrongs to any persons.[107] In order to create a constructive crime based on one of these underlying regulatory crimes of strict liability, then, there must exist a wrong in addition to the commission of such a regulatory offence. By injecting a requirement of minimal fault, we move from conduct that is unconnected to wrongs to any persons (driving whilst uninsured) to conduct that wrongs others, by increasing the risk of harm on the roads.[108] In our view it is preferable that a court articulate these reasons, rather than obscuring them within causation, and adding a 'harm within the wrong' inquiry to strict liability offences facilitates this. We reach the same conclusion as the Supreme Court, but by using a principle that could have universal application throughout the criminal law.

The benefit of this 'harm within the wrong' approach that requires the identification of the purpose and scope of the rule prior to addressing causation is that it leaves the meaning of causation otherwise unaltered. If the Supreme Court's insistence that causation necessarily involves fault were to establish a general principle, it would change the outcome in a number of the earlier pollution cases. Adopting a 'harm within the

105 A Summers, 'Common-Sense Causation in Law' (2019) 39 Oxford Journal of Legal Studies 50–73.

106 See R A Duff, 'Rule-Violations and Wrongdoings' in S Shute and A P Simester, *Criminal Law Theory: Doctrines of the General Part* (Oxford University Press 2002) 55.

107 R A Duff 'Wrongs and Responsibility for Wrongs' in M Dyson (ed), *Unravelling Tort and Crime* (n 2) 98. Elsewhere Duff distinguishes between harmful wrongs and wrongful harms, noting that speeding is an implicit endangerment offence and carrying a firearm is an indirect endangerment offence, which suggests that these offences are neither harmful wrongs nor wrongful harms without some further wrongful action on the part of the agent: R A Duff, 'Criminalizing Endangerment' in R A Duff and S Green, *Defining Crimes: Essays on the Special Part of the Criminal Law* (Oxford University Press 2005).

108 It might be argued that, if the wrong lies in the creation of a second-order harm, or risk, that we should retain the label of 'harm within the risk', as used in tort. Our reasoning, though, is that the purpose of the criminal law, as distinct from tort, is to censure wrongdoing, and it is therefore essential to focus on the wrong that is being censured.

wrong' test would facilitate confining the scope of *Hughes* to the s 3ZB offence and its ilk,[109] forcing the courts to deal with the issue on a case-by-case basis, rather than providing a general principle of causation. Under 'harm within the wrong', the environmental cases would then remain unchanged, as is clear if we revisit *Alphacell v Woodward*.[110] The defendant in that case was found guilty of the offence under s 2(1)(a) Rivers (Prevention of Pollution) Act 1951 establishing that D commits an offence 'if he causes or knowingly permits to enter a stream any poisonous, noxious or polluting matter'. Although it was not established that there was any negligence on D's part, the pump in their water tank had become blocked by branches and the overflowing pollutant drained into a stream. They were still held to have 'caused' the pollutant to enter the stream since:

> . . . the whole complex operation which might lead to this result was an operation deliberately conducted by the appellants and I fail to see how a defect in one stage of it, even if we can assume that this happened without their negligence, can enable them to say they did not cause the pollution.[111]

Lord Salmon explicitly drew the distinction between causal and legal responsibility, stating 'giving the word "cause" its ordinary and natural meaning, anyone may cause something to happen intentionally or negligently or inadvertently without negligence and without intention'.[112] If we were to adopt the Supreme Court's definition of causation in *Hughes*, then we would require some element of fault in a pollution case such as *Alphacell*, whereas the 'harm within the wrong' approach allows us to identify that this is a risk-based strict liability offence such that the conduct to which we apply the test of causation is the act of engaging in activity with pollutants at all.

The key to the controversy surrounding s 3ZB is the same as that identified by Padfield in her illuminating article in which she compares the approach taken to causation in the pollution case of *National Rivers Authority v Yorkshire Water Services Ltd*[113] to the approach taken in cases of homicide. The problem, she suggested, is that 'we are reluctant to convict of homicide someone whom we do not blame'.[114] Padfield wrote the article long before a homicide offence constructed on a regulatory driving offence had been created, as it now has been under the s 3ZB offence. What the pollution offences have in common with s 3ZB, however, is that they are result crimes, despite also being crimes of strict liability. In relation to traditional offences requiring *mens rea*, the policy questions which dictate determination of fair imputation can be dealt with as part of the *mens rea* of the offence. As noted by Hart and Honoré, if a court 'is satisfied that it has before it on a charge of murder a criminal who wilfully inflicted grave injury, without which death would not have occurred', the pertinent question is 'whether it [is] socially advantageous to give legal effect to the relation between the defendant's acts and the death'. But what of offences which dispense with the need for *mens rea* for the purposes of efficiency? In environmental cases, such as *Empress*, *Alphacell* and *Yorkshire Water Services Ltd*, it is arguable that efficiency outweighs the absence of blame; it is socially advantageous to give effect to the relation between the defendant's acts (of e.g. discharging sewage into a river) and the result by holding that D caused the harm, given that the harm is within the wrong

109 Such as the aggravated TWOC (taking without consent) offence under s 12A Theft Act, as applied in *R v Taylor* [2016] UKSC 5.

110 *Alphacell Ltd v Woodward* [1972] AC 824 (HL).

111 Ibid 834 (Lord Wilberforce).

112 Ibid 847 (Lord Salmon).

113 [1995] 1 All ER 225.

114 Padfield (n 74) 691.

which the strict liability offence was designed to prevent (pollution). The individual or company convicted is censured for failing to ensure that a river is not polluted by the activities from which it profits.

Other strict liability regulatory offences do not raise issues of causation because they tend to be *conduct* crimes not requiring any particular harm to have been caused, such as the offences under the Health and Safety at Work Act 1974 which punish the second-order harm of creating a risk of death or injury. Where that risk materialises and death is caused, the offence of corporate manslaughter is a *result* crime, where the 'harm within the wrong' requirement might be put to the test, with the prosecution having to prove that a gross breach of a duty of care *caused* V's death. In determining whether there was a gross breach of duty, the jury will have regard to any failure to comply with health and safety legislation.[115] It is in this way that the 'harm within the wrong' test might be applied to corporate manslaughter: the wrong in failing to discharge the duty under s 2 Health and Safety at Work Act 1974, for example, is the creation of a risk to the health, safety and welfare at work of an employer's employees. Where that risk materialises in death, it is clear that the harm is located within the wrong that legislation was seeking to prevent and, provided the other elements of the corporate manslaughter offence are established (including that the breach was *gross*), liability for causing death should ensue.

4 Concluding remarks

The fact that a case on the requirements under s 3ZB came to the attention of the Supreme Court so swiftly is to be welcomed, given the confusion amongst lawyers tasked with its application.[116] Some prosecutors had previously suggested that the use of the word 'causing' under s 3ZB was problematic and incongruent with the purpose of the law, suggesting that it would have been better to have used the terminology 'owing to the presence of a vehicle on the road'[117] (the terminology used under s 12A Theft Act) if the government really did think it appropriate to introduce a strict liability vehicular homicide offence. Yet, as the case of *Taylor* shows, with the Supreme Court now having interpreted s 12A Theft Act in light of *Hughes*, it was not the specific choice of words in the statutory provision that was in issue; rather that the courts are objecting to strict liability homicide offences at all. This is made clear by Lord Sumption in *Taylor*, stating that the 'fundamental reason why the appeal should succeed' is that to do otherwise would be to accept the Crown's invitation to treat s 2A as imposing strict liability for the aggravating factors.[118]

This paper has identified the complexity of the factors that shape the approach to causation in criminal law and negligence, showing that causation rules depend not only on the aims and justifications of the particular branch of law, but also the all or nothing nature of criminal liability compared to the allocation of responsibility that is possible within negligence and, crucially, on the place causation occupies in relation to the other elements of liability. Strict liability, whether in tort law or criminal law, raises distinct issues and is imposed for different reasons, and the introduction of a 'harm within the wrong' test for strict liability result crimes may help ensure that legal causation can be fully addressed in a way that reflects the functions of strict liability while confining liability within appropriate bounds. In an ideal world the legislature would respect the

115 Corporate Manslaughter and Corporate Homicide Act 2007, s 8.

116 See S Kyd Cunningham, 'Has Law Reform Policy been Driven in the Right Direction? How the New Causing Death by Driving Offences are Operating in Practice' (2013) Criminal Law Review 711, 720–721.

117 Ibid.

118 *R v Taylor* (n 54) [23].

principle of correspondence to prevent creation of inappropriate strict liability homicide offences in the first place. However, where they do come into existence, the courts are right to develop a principle that avoids injustice prevailing. The Supreme Court's approach in *Hughes* and *Taylor* has the potential to develop into a 'harm within the wrong' test to ensure that requirements of causation can give effect to the role of the criminal law in censuring the blameworthy.

NILQ 70(4): 445–465

Charitable purposes, demonstrable benefit and the role of the Charity Commission: the fourth *Pemsel* head recast

LARA MCMURTRY*

Keele University

Abstract

The Charities Act 2011 requires that charitable purposes must fit within one or more of the statutory descriptions of 'charity' and, demonstrably, be in the public benefit. From the perspective of the now dismantled fourth Pemsel *head of charity, this article examines the elusive concept of a charitable purpose and the misconceived statutory public benefit requirement. The guidance and decision-making roles of the Charity Commission are appraised and problems in applying the current law exposed. As few charity cases reach the courts or Upper Tribunal, this work analyses the limited, but important, registration appeals that have reached the First-tier Tribunal. It will explore the emerging trends in charity appeals, the divergent approaches adopted by the Commission and the Tribunal and the lamentable degree of unpredictability facing prospective charity trustees. It concludes that an absence of clear-thinking pervades this vital aspect of charity law and invites reworking and revision.*

Keywords: charity; charitable purpose; Charity Commission; Charity Tribunal; public benefit

Introduction

Under the fourfold classification of charity formulated by Lord Macnaghten in *Commissioners for Special Purposes of the Income Tax v Pemsel*,[1] the final head focused upon 'other purposes beneficial to the community'. Until its demise brought about by the Charities Act 2006,[2] this discrete category offered a repository for the diverse grouping of purposes deemed charitable at law and outside the other heads of relieving poverty, advancing education or advancing religion.[3] It comprised a sundry grouping of established sub-categories, such as the promotion of health,[4] the welfare of animals[5] and the promotion of public works, services or amenities.[6] Taking its inspiration from the Preamble to the 1601 Statute of Charitable Uses, which contained a non-exhaustive list of purposes

* I am grateful to Professor Tony Bradney and Professor Warren Barr for their comments on an earlier draft.

1 [1891] AC 531.

2 Repealed by the Charities Act 2011. Unless otherwise stated, all statutory references are to the 2011 Act.

3 To be charitable, a trust must be for a purpose deemed charitable at law, exclusively charitable and for the public benefit.

4 *Re Hillier* [1944] 1 All ER 480.

5 *Re Grove-Grady* [1929] 1 Ch 557.

6 *British Museum Trustees v White* (1826) 2 Sim & St 594.

deemed charitable at the time of that enactment, the fourth head represented a body of law of 'intolerable messiness',[7] built up, 'not logically but empirically'.[8] Nevertheless, it permitted the inclusion of novel purposes by a process of analogy[9] and, thereby, maintained flexibility in the law.[10] This latter characteristic was appraised as, 'its greatest strength and most valuable feature'.[11]

The fourth head drew upon centuries of accumulated, and sometimes discordant, case law developments, traceable back to the 1601 Preamble. The Preamble had modestly attempted to illustrate the purposes that might be charitable rather than offering any definition of what comprises a charitable purpose.[12] It mentioned, for example, the 'relief of aged, impotent and poor people', the 'maintenance of sick and maimed soldiers' and the 'redemption of prisoners'. Nevertheless, it was to assume major significance in the development of charity law, which was to recognise only that 'which the Statute enumerates or which by analogies are within its spirit and intendment'.[13] Despite its formal interment by the Charities Act 1960, judicial precedent ensured that the Preamble continued to exert influence, especially so with the fourth *Pemsel* head. Indeed, the fourth head remained inextricably bound to the archaic requirement that any newly admitted purpose must be within the equity of the statute.[14] If not, the traditional wisdom was that, however beneficial, the purpose could not be charitable at law.[15]

The requirement that a novel purpose was to be sited within the spirit and intendment of the Preamble encouraged the courts to be creative and sometimes stretch analogies, 'almost to breaking point'.[16] The most striking example was provided in *Vancouver Regional Freenet Association v Minister of National Revenue* where the provision of internet access was held to be analogous with the repair of highways.[17] Unsurprisingly, the case for reform was urgent and compelling. As Alan Milburn admitted, 'today's charitable sector would be completely unrecognisable to those who set out the first framework for charitable law, and it is characterised by its huge diversity and enormous size'.[18] The Charities Act 2011 marked the culmination of a review and consultation process aimed at reinventing the legal definition of charity with heightened emphasis upon public benefit.[19] This delivered the fatal blow to the *Pemsel* classification, which was replaced by a more comprehensive list of statutory descriptors. The fourth head was dismantled and reconstituted in the

7 H Kidd, 'Definition of Charity: A Modest Proposal for a Bit of Divorce' (1992–1993) 1(1) Charity Law and Practice Review 5, 6.

8 See *Oppenheim v Tobacco Securities Trust Co Ltd* [1951] AC 297, 309 per Lord Normand.

9 See S*cottish Burial Reform and Cremation Society v Glasgow City Corp* [1968] AC 138.

10 See *Re Foveaux* [1895] 2 Ch 501. As will become clear, the development of the law by analogy remains integral to charity law.

11 See Home Office, *Charities – A Framework for the Future* (Cm 694, 1989) [2.11].

12 See B Bromley, '1601 Preamble: The State's Agenda for Charity' (2002) 7(3) Charity Law and Practice Review 209.

13 *Morice v Bishop of Durham* (1804) 9 Ves 399, 405 per Sir William Grant MR.

14 A continuing process described by Kidd (n 7) as that, 'game of inspecting the Sibylline book of 1601'.

15 *Re Strakosch* [1949] Ch 529; see also Lord Simonds in *Gilmour v Coats* [1949] AC 426, 442.

16 *Scottish Burial Reform and Cremation Society* (n 9) 153 per Lord Upjohn.

17 (1996) 137 DLR 406. See further A Iwobi, 'Rolling Down the Information Highway in Search of Charitable Status' (1998) 5 Web Journal of Current Legal Issues.

18 HC Deb 26 June 2006, vol 448, col 44.

19 See, generally, the Labour Government's Strategy Unit Report, *Private Action, Public Benefit: A Review of Charities and the Wider Not-for-Profit Sector* (Cabinet Office, September 2002).

guise of the individual descriptions set out in s 3(1)(d)–(l).[20] The laudable ambition was that this 'expanded list of purposes would make the overall framework much clearer both for charities and for the public'.[21] Accordingly, Parliament required that a purpose must fall within one or more of the 13 replacement statutory heads and that it be demonstrably in the public benefit.

Regrettably, the Act refrained from charting the parameters of its public benefit requirement. Two related provisions, however, assume relevance. First, s 4(3) states that public benefit is to be interpreted 'as that term is understood' for the purposes of the law of charity in England and Wales. This reinforces the pre-existing understanding of public benefit[22] and ensures that the pre-existing case law, however anomalous, must necessarily pervade the current law.[23] Secondly, s 4(2) has sparked controversy, despite its seemingly innocuous wording. This provides that, in determining whether the public benefit requirement has been satisfied, it is 'not to be presumed that a purpose of a particular description is for the public benefit'. As will be discussed, this provision was designed to end the operation of the so-called 'presumption' of public benefit[24] that was generally, but not universally, believed to have aided the courts in previous times.[25] Instead, an intrinsic benefit must now be demonstrated affirmatively in all cases. A key problem that is addressed in this article concerns how this aspect of public benefit is to be evidenced in any given case, particularly those under the unpacked fourth head. The enigmatic nature of public benefit continues to cause problems both for prospective applicants and the Charity Commission, and it may now be time to consider afresh the prospect of formulating a statutory definition. Parliament's failure to draw out meaning, and to deal with attendant implications, has generated unnecessary and unwanted practical difficulties. Accordingly, the present work will appraise the difficulties of establishing a public benefit within this changed legal landscape as dominated by the Charity Commission and the Charity Tribunal, blighted by conceptual uncertainty and obscured by linguistic unfamiliarity.

The decision-makers

Taking centre stage in the post-legislative arena is the Charity Commission, which determines which institutions and purposes are to be categorised as charitable.[26] The importance of the registration process cannot be underestimated as it provides an assurance that the extensive fiscal privileges that attach to charitable status are deserved. There are also the reputational advantages that come with enhanced public standing, instilling confidence that such charities operate for the public benefit. Significantly, the Commission is, by s 17(1), tasked to provide guidance to promote awareness and understanding of the public benefit requirement. While the Commission is not technically

20 Some of the newly listed purposes were imported from the other pre-existing heads, for example, the promotion of the arts (in s 3(1)(d)) from the 'advancement of education'.

21 *Private Action, Public Benefit* (n 19) [4.12].

22 As J Hackney observed in 'Charities and Public Benefit' (2008) Law Quarterly Review 347, 347, 'this section effects no change at all'.

23 See, for example, the strikingly conservative approach of the Upper Tribunal in *AG v Charity Commission* [2012] UKUT 420 (TCC), upholding the charitable status of benevolent funds for the relief of poverty amongst family members and employees.

24 See Hillary Armstrong, HC Deb 26 June 2006, vol 448, col 25.

25 The existence of this presumption was denied in *ISC v Charity Commission* [2011] UKUT 421 (TCC) (hereafter *ISC*) but previously was thought to apply to trusts in the first three *Pemsel* categories.

26 See generally R Meakin, *The Law of Charitable Status: Maintenance and Removal* (Cambridge University Press 2008).

a lawmaker, it undoubtedly enjoys a role as a quasi-judicial arbiter of status. Absent rectification of the register, s 37(1) states that registration by the Commission is conclusive of charitable status. It is important, therefore, that registration decisions are robust, fair and consistent and that the public benefit guidance is comprehensive and on all fours with traditional judicial wisdom. This article will establish that these ideals are not always achieved. The Commission has already suffered the chastening effect of a successful legal challenge to its public benefit guidance as regards education[27] with the subsequent need to revisit its guidance and accompanying legal analysis.[28] Admittedly, the Commission is not helped in its task by the paucity of cases reaching the courts or the Upper Tribunal. Only a modest measure of oversight and redress is provided by the First-tier Tribunal (Charity).[29] Although not a court of precedent and with its proceedings being geared to 'the procedural informality and quasi-inquisitorial approach characteristic of tribunals',[30] its key role is to consider afresh contested decisions as to eligibility for charity registration. As will become clear, the Commission's ideals are not always shared and important differences in approach between the bodies have come to the fore.

A limited (but ever-growing) number of First-tier Tribunal decisions paint a revealing picture of the existing tension and uncertainty in the Commission's explication and application of the revamped code. From the perspective of the fourth head in its contemporary and disassembled form, this article will critically examine the interpretation and impact of this redefinition of charitable purpose and public benefit. This is an aspect of charity law that has hitherto been neglected, with academic and judicial attention having been focused instead upon the more conspicuous heads of poverty, education and religion. Two problems inherent in the redefined concept of a charitable purpose will be exposed and examined. The first is conceptual and derives from the lack of settled meaning in the statutory language and open-endedness of the descriptions. This is directly attributable to the failure of Parliament to offer adequate statutory guidance or clarification. The second lies in the climate of uncertainty that shrouds public benefit, in particular, the Commission's interpretation of benefit and what it asks of an organisation in order that its public 'benefit' be established. It will become apparent that the Commission operates with an acute awareness of the threat of ambiguity in an organisation's objects and, by insisting upon compelling formal proof of impact, sets an unrealistically high bar to be overcome. While the desire to impose rigour and order in its decision-making process may be laudable, this article will demonstrate that the Commission becomes overly mired in legal technicality. Unsurprisingly, this tendency impedes the valuable, innovative and philanthropic efforts of those at the forefront of charitable endeavour and generates unpredictability and uncertainty for prospective charity trustees. Both these problems will now be considered in turn.

Inherent uncertainty and conceptual misgivings

The decision to update the list of charitable purposes was popular, uncontested and produced a wider range of descriptions of charity with a decidedly contemporary feel. With some notable amendments in scope, the first three heads were retained and duly

27 *ISC* (n 25).

28 See, generally, M Synge, *The New Public Benefit Requirement* (Hart 2015) 219–231.

29 See D Morris, 'The First-tier Tribunal (Charity): Enhanced Access to Justice for Charities or a Case of David versus Goliath' [2010] Civil Justice Quarterly 491.

30 A McKenna, 'Should the Charity Tribunal be Reformed' (2011–2012) 14 Charity Law and Practice Review 1, 4.

modified.[31] As mentioned earlier, the unhelpfully opaque 'other purposes' head was unceremoniously dismantled and replaced by specific, individual descriptions of the purposes. Undoubtedly, the revised list of statutory descriptions reflects more completely the breadth and scope of charitable purposes pursued in modern times. While a number of these purposes had long been held to be charitable,[32] others are of a more recent vintage.[33] Importantly, the promotion of existing charitable purposes to new delineated statutory categories serves to enhance the status of those purposes and emphasise their legitimacy. It also underscores an equivalence of treatment with no evident intention to prioritise or distinguish on the merits of a claim to fiscal privileges.[34] Hence, community-orientated ends, such as the promotion of amateur sport, sit alongside fundamental humanitarian purposes such as the promotion of human rights or the saving of lives. The forms of benefit are diverse, often focused upon fulfilling a specific need,[35] but sometimes are indirect and of broader benefit to the wider community.[36] In the absence of demand to extend its number,[37] there is a sense of permanence to this statutory list. It is not possible to alter the descriptions in light of changing social circumstances or with a view to adding to, or subtracting from, the list. The flexibility so prized at common law is, however, preserved by s 3(1)(m). This residual category reflects, 'the way in which charity law has developed in extending incrementally . . . analogy upon analogy, the legal concept of charity'.[38] Recognised purposes that do not feature in the list continue to be charitable,[39] as do purposes that 'may reasonably be regarded as analogous to, or within the spirit of' a recognised charitable purpose or a purpose analogous thereto.[40]

There is no doubt that the shift from the historically recognised heads of charity to a longer list of statutory descriptions has generated concerns for organisations whose purposes might have formerly fallen under the fourth *Pemsel* head. As Buckley recognised, 'a lot of unanswered questions remain . . . it is both uncertain what these new categories comprise and what, in turn, should be the appropriate test of public benefit in respect of each of them'.[41] Obvious problems are faced by prospective charity trustees who are confronted with a list of descriptions that are couched in novel language and are not defined more fully. As will become clear, the stark reality is that an institution has to demonstrate that it is established exclusively for charitable purposes in circumstances where the confines of the descriptions are not fully understood. Hence, the key

31 The relief of poverty was widened to encompass the 'prevention or relief of poverty' (s 3(1)(a)), and the scope of religion was to include 'a religion which involves belief in more than one god' or that 'does not involve belief in a god' (s 3(2)(a)).

32 An obvious example is the advancement of the arts, which appears explicitly in s 3(1)(d) and is recognised as charitable in celebrated cases such as *Re Verrall* [1916] 1 Ch 100 (National Trust).

33 For example, the promotion of amateur sport, which now enjoys statutory recognition in s 3(1)(g).

34 For relevant discussions of charity and fiscal privilege, see S Bright, 'Charity and Trusts for the Public Benefit – Time for a Re-think' [1989] Conveyancer and Property Lawyer 28; J Warburton, '"Charity" – One Definition for all Tax Purposes in the New Millennium?' [2000] BTR 144.

35 See s 3(1)(j) ('the relief of those in need by reason of youth, age, ill-health ...').

36 For example, the advancement of environmental protection or improvement (s 3(1)(i)).

37 See Lord Hodgson, *Trusted and Independent: Giving Charity Back to Charities – Review of the Charities Act 2006* (TSO, July 2012) [4.15].

38 *ISC* (n 25) [77].

39 As H Picarda observed in *The Law and Practice Relating to Modern Charities* (Butterworths 1977) 108: 'No list of charitable purposes can be exhaustive and it is perhaps inevitable that any classification of purposes will leave a rag bag of individual charitable gifts with no easily discernible common thread.'

40 S 3(1)(m)(ii), (iii).

41 C Buckley, 'The Charities Act 2006: Consolidation or Reform?' (2008) 11(1) Charity Law and Practice Review 1, 42.

registration decisions of the Commission, coupled with the trickle of appeals to the First-tier Tribunal, offer for the first time the opportunity to examine this emerging obstacle to charitable status.[42]

Decoding the statutory language

Undoubtedly, the unfamiliarity of the statutory wording and the open-endedness of the descriptions combine to generate conceptual and practical problems for the Commission. In contrast with the well-established categories of charity, many of the new descriptions lack both clarity and meaning. The finding that a particular purpose falls within the concept of education, for example, renders it self-evidently beneficial because the legal meaning of education is so innately understood as including certain elements and precluding others. It necessarily excludes a range of purposes that might be deemed political, subversive, lacking in utility or subjective merit. As Garton explains, 'a prima facie charitable purpose is rendered non-charitable because it delivers only an intangible or inefficacious benefit, or has detrimental side effects'.[43] In comparison, many of the newly formed descriptions remain vague with untested boundaries and only limited or fact-specific case law illustrations from which to distil general principles. Particular problems for prospective trustees have emerged in the context of human rights and community development.

In *Human Dignity Trust v Charity Commission for England and Wales*,[44] the Human Dignity Trust (HDT) was established to assist persons whose human rights were violated by the criminalisation of private, adult, consensual, homosexual conduct, and it sought recognition as a charity under s 3(1)(h) of the 2011 Act. The Commission refused initially to register the Trust because, in campaigning for a change in the law, it was political in nature. Of present interest, however, is the Commission's strikingly conservative view of the concept of human rights, surprisingly at odds with its own prior thinking. Citing the lack of clarity as to what human rights means in English law, the Commission favoured a narrow interpretation of the concept, which was to the detriment of the Trust. Ironically, this contradicted its previous, valuable efforts to produce encompassing non-statutory guidance with a broad conception of human rights,[45] upon which the Trust had not unreasonably sought to rely. It fell to the First-tier Tribunal to take the more confident stand that the 'broad, undefined description of a purpose in the Act was intended to build upon the valuable work already undertaken by the Charity Commission'[46] and that the scope of rights falling within the description was apt to, 'evolve and change from time to time'.[47] With characteristic liberality of reasoning,[48] the appeal against the refusal of the Commission to register the Trust as a charity was allowed. Concerned with interpreting and protecting superior constitutional rights, the purposes were, in its view, geared to upholding the law, not changing it.

42 Note that First-tier Tribunal decisions, 'turn on their own facts and have no precedent value': A McKenna, 'The Charity Tribunal – Where To and From' (2014) 4 Private Client Business 213, 214.

43 J Garton, *Public Benefit in Charity Law* (Oxford University Press 2013) 47.

44 CA/2013/0013, 9 July 2014 (FTT (Charity)) (hereafter *HDT*).

45 Charity Commission, *The Promotion of Human Rights (RR12)* (1 January 2012).

46 *HDT* (n 44) [42].

47 Ibid [44].

48 Here, the Tribunal has been criticised for its reliance upon, 'highly vague generalities to describe what "human rights" means for purposes of charity law' (see A Parachin, 'Charity, Politics and Neutrality' (2015–2016) 18 Charity Law and Practice Review 23, 37).

A further illustration of uncertainty may be found in the context of 'community development', which forms part of s 3(1)(e) along with the advancement of citizenship.[49] It is understandable that 'community development' should assume a place on the statutory list. There is a wealth of Commission guidance on what might be caught by this description,[50] and it is largely due to the Commission's endeavours that community development has become a category of modern-day charitable enterprise. The Commission's decision in 2003 to register Guidestar UK, for example, admitted as charitable the new purpose of promoting the voluntary sector.[51] It is logical, moreover, that trustees might look to that existing Commission guidance for assistance in interpreting the law. The problem remains, however, with the paucity of assistance offered to trustees, anxious to demonstrate that their purposes are exclusively charitable within the bounds of the description. Such was evident in *Full Fact v Charity Commission for England and Wales*,[52] where the independent fact-checking organisation, Full Fact, failed in its initial attempts to obtain charitable status with purposes that included the advancement of civic responsibility and engagement. The First-tier Tribunal admitted that the 'promotion of civic responsibility' was, 'undefined by the courts'.[53] It was, moreover, troubled that the term 'civic engagement' was not necessarily within the same sphere as 'civic responsibility' and, accordingly: 'It cannot be assumed that "engagement" is always positive and capable of providing benefit.'[54] Although the Commission was later to accept a renewed application from Full Fact, it was with an objects clause that fell entirely within s 3(1)(b) of the Act (i.e. the advancement of education).[55] It is troubling that Full Fact was unable to establish that its purposes fell within the specific sphere of charity it preferred to pursue. It came to appreciate that it was manifestly more difficult to structure its objects to those desired ends.

Exclusivity of charitable purpose

Not only is the statutory language novel, the new descriptions are troublingly open-ended, such that both charitable and non-charitable activities might be pursued.[56] For Lord Hodgson, it was contentious that a statutory list of charitable purposes 'may include purposes which are not in fact charitable'.[57] He readily identified an anomaly that 'an organisation with the express object of "relieving poverty" may be regarded by the Commission as having charitable purposes . . . but an organisation with the express object of "advancing amateur sport" may not'.[58] Such is well illustrated in *Uturn* where the company lost its appeal against the refusal by the Commission to register it as a charity. The primary object of Uturn was to promote and set up street associations to 'advance citizenship and community development'. More specifically, its ambition was 'to bring the

49 S 3(2)(c)(ii) provides that this includes the promotion of civic responsibility and volunteering.

50 See Charity Commission, *The Promotion of Rural and Urban Regeneration (RR2)* (1 March 1999); *The Promotion of Community Capacity Building (RR5)* (1 November 2000).

51 See Charity Commission, *The Promotion of the Voluntary Sector for the Benefit of the Public (RR13)* (1 September 2004).

52 CA/2011/0001, 26 July 2011 (FTT (Charity)) (hereafter *Full Fact*).

53 Ibid [7.3].

54 Ibid.

55 17 September 2014.

56 See *Uturn UK CIC v Charity Commission for England and Wales* CA/2011/0006, 27 February 2012 (FTT (Charity)) (hereafter *Uturn*) where at [8] the 12 specific heads of purpose were seen as no more than, 'potentially charitable'.

57 Hodgson (n 37) [4.15.2].

58 Ibid.

residents of streets together in local groupings with a framework that will engender civic responsibility and volunteering'. This primary purpose appeared to fall squarely within s 3(2)(c), with street associations the particular vehicle adopted to promote one of the stated descriptions of charity. The Tribunal, however, shared the concern of the Commission that the expression 'street association' did not carry any settled meaning in law and warned that 'the framing of a purpose by reference to an expression, the meaning of which is uncertain, risks the meaning of the purpose itself being unclear or uncertain'.[59] In deciding that the purposes of the company were not exclusively charitable, the Tribunal made clear that, 'for any purpose to be charitable (whether it is new or well established), it must be sufficiently certain, to permit an assessment of whether it falls exclusively within the heads of purpose listed . . . in the Act'.[60] This is a conclusion manifestly more difficult to reach where the content of those heads of purpose remains alarmingly unsettled.

The decision in *Uturn* is a reminder that merely to adopt the language of the descriptions as a means to communicate the objects of an organisation will not necessarily satisfy the Commission or the Tribunal that an organisation is confined to the pursuit of exclusively charitable purposes. More is required than a statement that a legally recognised benefit lies at the heart of an organisation's purposes.[61] By corollary, it can be argued that more might be expected in terms of effective guidance to prospective charity trustees seeking to situate their objects within the new descriptions. In *Uturn*, the Commission identified a lack of clarity regarding the operation of street associations. The admitted concern (shared also by the Tribunal) was that an association might 'provide services to individuals which confer a private benefit'. The Commission was looking, therefore, for a high degree of certainty regarding the future operational activities of each association that would ensure that, in practice, it worked towards exclusively charitable ends. Uturn could not meet that threshold even though its street association project offered a fresh and innovative response to a pressing social problem.

Ambiguity and artifice

As considered above, the lack of settled meaning underpinning the revised descriptions has resulted in a problematic degree of conceptual and legal uncertainty. In consequence, the Charity Commission feels obliged to look beyond the institution's governing document and to consider extrinsic evidence. This approach is not without controversy. In determining whether or not an organisation is established for charitable purposes only, the initial task is to identify the 'particular purpose' (or purposes) of the institution.[62] This involves ascertaining, 'what the institution was set up to do . . . not how it would achieve its objects or whether its subsequent activities are in accordance'.[63] The authorities are clear that the initial focus must be on the objects clause or its equivalent in the institution's governing document.[64] The relevant test here is not an activities test, but is instead a test of what an organisation is established to do. As Synge explains, 'charitable status is not a transient quality which depends on a charity's activities, but a near permanent quality

59 *Uturn* (n 56) [20].

60 Ibid [27].

61 A point emphasised in *Crocels Community Media Group v Charity Commission for England and Wales* CA/2015/0009, 18 July 2016 (FTT (Charity) [23].

62 *ISC* (n 25) [82].

63 Ibid [188].

64 See *Helena Partnerships Ltd v Revenue and Customs Commissioners* [2011] UKUT 271.

which depends on the purposes for which an institution is established'.[65] This echoes the sentiment expressed by Buckley LJ that it is irrelevant 'to inquire what the motives of the founders were, or how they contemplated or intended that the . . . [organisation] should operate, or how it has in fact operated'.[66]

Nevertheless, in *Helena Partnerships Ltd v Revenue and Customs Commissioners*, the Upper Tribunal acknowledged that the court may examine activities 'where there is a doubt or ambiguity about whether the objects of an institution are charitable'.[67] This qualification to the general rule appears to distinguish between the process of working out what the particular purposes of an organisation are (i.e. what it is established to do) and the attendant matter of whether those purposes are charitable purposes. In *Incorporated Council of Law Reporting for England and Wales v AG*, Buckley LJ offered support for that distinction. He admitted that 'in order to determine whether an object, the scope of which has been ascertained by due processes of construction, is a charitable purpose it may be necessary to have regard to evidence to discover the consequences of pursuing that object'.[68] There the purpose was identified as being the preparation and publication of law reports and the court felt it necessary to consider extrinsic evidence as to whether the performance of its objects achieved an exclusively charitable purpose.

Purposes or activities?

Although the Commission purports to follow the courts in applying a purposes test, it readily employs an analysis of activities and proposed activities to gain a proper understanding of the body's true purpose. This enquiry is separate to the determination of whether those purposes will furnish a public benefit,[69] and it involves an additional, unjustified measure of scrutiny for prospective charities. For Luxton, this approach is problematic because it ignores 'a vital difference between a legitimate regard to activities in order to determine whether they are consistent with the body's charitable purposes and an improper attempt to require charities to show public benefit in their activities'.[70] Nevertheless, the Commission continues to undertake this analysis and does so readily when the purposes of an organisation appear to it to be unclear or ambiguous. Indeed, where the purposes of organisations fall within the new descriptions, this is the prima facie starting point. In *HDT*, for example, the primary concern was the political element to the purposes, but the Commission was explicit that where an institution purports to advance human rights, it would:

> . . . generally be necessary to consider extrinsic evidence (including evidence about the nature of its activities or proposed activities) because, firstly, there is as yet no 'particular meaning' in charity law of the term 'human rights' and secondly, the promotion of human rights is a broad concept which could include non-charitable activities.[71]

65 M Synge, 'A State of Flux in Public Benefit across the UK, Ireland and Europe' (2013–2014) 16 Charity Law and Practice Review 163, 168.

66 *Incorporated Society of Law Reporting for England and Wales v A-G* [1972] Ch 73, 99.

67 *Helena Partnerships* (n 64) [20]. This accords with the views expressed by Garton that there are ways in which activities 'could *potentially* play a role' (original emphasis) such as in the case of a novel purpose, to 'determine whether that purpose falls within the categories of charity': J Garton, 'Charitable Purposes and Activities' (2014) 67 Current Legal Problems 373, 389.

68 [1972] Ch 73, 99.

69 See, for example, the decision of the Commission not to register the Countryside Alliance: <www.gov.uk/government/publications/the-countryside-alliance> (23 March 2017) [16].

70 P Luxton, *Making Law? Parliament v The Charity Commission* (Politeia 2009) 18.

71 HDT (n 44).

Tellingly, the First-tier Tribunal did not share that hypothesis, finding that HDT's purposes were clear without recourse to extrinsic evidence. The Commission, by contrast, was entangled in an analysis of activities that led it to conclude that a parallel purpose lurked beneath and that HDT was not established for exclusively charitable purposes.

In V*ernor-Miles v Charity Commission for England and Wales*, moreover, the stated purpose of the Independent Press Regulation Trust (IPRT), was to 'promote for the public benefit high standards of ethical conduct and best practice in journalism'.[72] A separate sub-clause contained a power to expand the objects clause, which included the provision of financial assistance towards the establishment and support of an independent press regulator. The Commission was quick to identify those purposes as ambiguous and focused instead upon the 'express means by which the purpose is to be furthered'. As a Leveson-compliant press regulator was yet to be established, the Commission felt that the purposes of IPRT were too vague and uncertain to conclude that they were exclusively charitable. In its view, IPRT was promoting a concept that had yet to be established as beneficial for the purposes of charity law, and it was unprepared to reason by analogy in its favour. Once again, the First-tier Tribunal did not agree that the purposes were unclear or ambiguous. It did not share the view of the Commission that a collateral purpose could be discerned and concluded that the particular purpose of IPRT was to be found in its objects clause alone.[73] It could find no evidence that a 'non-ancillary private benefit would accrue to those in the media industry' and highlighted the 'unchallenged evidence' that any private benefit would be incidental.[74] It was wrong in principle to require the appellants to prove a negative, which was the insurmountable burden the Commission had imposed on IPRT.

Ancillary purposes

Analysis of the registration decisions reveals a divergence of approach to the key matters of determining an institution's objects and distinguishing purposes from the means of their pursuit. Indeed, the Commission might seem somewhat over-zealous in its efforts to unearth collateral purposes that it suspects might be masked or hidden by the applicant. Judicial support for this robust type of investigative inquiry is, nonetheless, discernible, particularly where political purposes are concerned. For example, *Southwood v AG*[75] advocated the use of extrinsic evidence to determine the 'true' (i.e. non-charitable) purposes of an institution and is frequently cited as authority for that proposition.[76] The briefing papers and project proposals available led the court to conclude that the trust was of an overtly political character and not exclusively charitable for the advancement of education. This has not been universally accepted as being the correct approach. Garton complains that 'The trust's purpose was reinterpreted . . . the correct approach should have been to declare the trust valid as being conceptually for the public benefit, although the activities carried on pursuant to it may well have been in breach of trust.'[77] Nevertheless, it is undeniable that the perceived ambiguity in the stated objects strongly encouraged the use of extrinsic evidence as an aid to interpretation.

72 CA/2014/0022 15 June 2015 (FTT (Charity)) (hereafter *IPRT*).

73 It benefited from a report outlining the public benefit of the Leveson recommendations (see below).

74 *IPRT* (n 72) [38].

75 [2000] WTLR 1199.

76 See, for example, P Luxton, *The Law of Charities* (Oxford University Press 2001) 440; P Luxton, 'A Three-Part Invention: Public Benefit under the Charity Commission' (2009) 11 Charity Law and Practice Review 19, 27.

77 Garton (n 43) 85.

Whatever the merits of a robust enquiry, the Commission has proved itself willing to move beyond even these limits and to seize upon the post-formation activities of a registered charity to make the case that it was not established for charitable purposes. Those problems were demonstrated in the litigation involving the Human Organ Preservation Research Trust (HOPRT).[78] This decision serves well to highlight some worrying features of the Commission's approach. HOPRT was established by deed of trust in 1990 and entered onto the Register in 1991. Its objects were overtly research-orientated. Primarily, it was to 'conduct promote develop and co-ordinate for the benefit of the public research' with a view to enabling the preservation of human organs after death for transplant purposes and to undertake research 'into the ageing process'. By 2017, the organisation was conducting research and training, providing information and advice, while catering also for the provision of cryopreservation services to members of the public and facilitating the transport of preserved bodies into storage. Following a period of heightened press attention[79] and spurred by a complaint by a member of the public, the Commission solicited evidence of HOPRT's post-formation activities and took the decision to remove it from the Register.[80] It did not reach its decision on the basis that HOPRT had ceased to operate as a charity. Instead, it felt that there was insufficient evidence to show that HOPRT had been established for exclusively charitable purposes. In the view of the Commission, its objects were and always had been the promotion and facilitation of cryopreservation, which did not fall within the purposes set out in its trust deed. Accordingly, its entry onto the Register had been a mistake requiring rectification.

On appeal, the First-tier Tribunal turned its attention to the purposes for which HOPRT was established, but could find no evidence of a sham, misunderstanding or miscommunication at the time of registration that might support the Commission's case of registration by mistake. On the contrary, on the evidence available,[81] the Tribunal found it 'more likely than not that it was the Commission who suggested the final wording of the objects clause or, at the very least, provided input on the words used'.[82] Furthermore, on a plain reading of the objects clause the Tribunal found no ambiguity. The objects as drafted were 'by design, generous and permissive of a wide range of research activities, including those which were not then feasible but which might become so in the future'.[83]

Although it expressed concern that the trust deed had 'become stale and in need of updating',[84] the Commission was reminded that the problem was a regulatory matter, unconnected with charitable status. Indeed, the focus of the appeal was not to 'conduct a review of all the possible ways in which a charity may satisfy the statutory test for registration'.[85] Instead, it was simply to determine what HOPRT was established to do.

78 *Hipkiss v Charity Commission for England and Wales* CA/2017/0014, 23 August 2018 (FTT (Charity)) (hereafter HOPRT).

79 Sparked by media coverage of *Re JS (A Child) (Disposal of Body: Prospective Orders)* [2017] 4 WLR 1 (a landmark decision allowing a terminally-ill, 14-year-old girl to preserve her body after death).

80 The Commission has the ability to do so under s 34(1)(a), (b).

81 With only limited contemporaneous documentary evidence, much of the evidence presented was circumstantial and hearsay.

82 *HOPRT* (n 78) [107].

83 Ibid [110]. Note that cryopreservation was not possible at the time the trust was executed.

84 Ibid [119].

85 Ibid [123].

The direction was that HOPRT be restored to the Register.[86] More broadly, however, the decision serves to expose the crippling uncertainties faced by organisations seeking to show that their purposes are clear and exclusively charitable from the objects clause in their governing instrument. The response of the Commission does not always help and its decision to act here was unpredictable, time-consuming and ultimately baseless.

A beneficial purpose?

As between the Commission and the Tribunal, the preceding analysis exposes the dichotomy in approach to discerning the purposes for which an institution is established and the matter of whether these purposes fall exclusively within one or more of the statutory descriptions. It is not sufficient, however, that a purpose falls within one or more of the new descriptions and that it is exclusively charitable. It must also be for the public benefit. The concept of public benefit is widely understood to have two interrelated aspects.[87] In its first sense, the purpose must be inherently beneficial in a way that the law deems charitable (i.e. the benefit element of public benefit). In its second sense, that benefit must be visited upon a sufficient section of the community (i.e. the public element of public benefit).[88] As this article is concerned only with public benefit in the first sense, it does not engage with the required reach or scope of that beneficial purpose once established. The task that remains is to analyse the approach of the Commission to the demonstration of benefit as part of the public benefit requirement. Again, the cases to date paint a picture of needless bureaucracy and conceptual conflation.

Isolating benefit

It is somewhat odd that a separate assessment of benefit may be required once a purpose has been shown to fall within one or more of the statutory descriptions.[89] The Strategy Unit Report emphasised that the purposes were 'chosen to represent the main areas of charitable activity that can reasonably be anticipated to continue to represent a public benefit'.[90] For the Upper Tribunal, the effect of the statutory wording is to separate out, in a highly artificial manner, the issue of whether the purpose of an institution falls within the descriptions of charity from whether it is beneficial so as to establish public 'benefit'.[91] The relevance of the statutory list of purposes is that it 'presents a hurdle: the purpose must fall within one or more of the categories and if it does not the question of public benefit is not relevant'.[92] The pertinent issue is 'not whether the categories are inherently or necessarily for the public benefit, the focus is on the particular purpose of a particular institution'.[93] Regrettably, it follows that the claimant with a particular purpose that falls within the descriptions must face further scrutiny with something additional to demonstrate. Albeit tentatively, Synge notes that a separate assessment of

86 The vulnerability of the beneficial class was clearly of influence as (ibid [102]) the Commission was rebuked for the failure to consider how best to support individuals in the event that HOPRT (see n 78) was removed from the Register.

87 ISC (n 25) [44].

88 J Jaconelli, 'Adjudicating on Charitable Status – A Reconsideration of the Elements' [2013] Conveyancer and Property Lawyer 96, 97, conceives of these elements as being qualitative and quantitative respectively.

89 As P Luxton (n 70) at 6 explains, 'the case law establishes that, once a purpose is shown to be charitable, it is necessarily for the public benefit in the first sense'.

90 *Private Action, Public Benefit* (n 19) [4.14].

91 *ISC* (n 25) [82].

92 Ibid.

93 Ibid.

benefit is not itself objectionable, but argues that 'inclusion on the list should be understood to signify that they [the purposes] are prima facie beneficial, so that satisfying the benefit aspect requires consideration of possible disqualifying factors, rather than separate evidence of benefit'.[94] This would seem to be an eminently sensible shift of emphasis, but it directly conflicts with the policy and practice of the Commission, which places the onus firmly on the applicant in every case to demonstrate beneficial impact.

Much attention has been devoted in an effort to discern the scope of the common law presumption of public benefit (if, indeed, it may properly be so described) and the effect of the Act in securing its jettison.[95] Beyond the enactment of this short statutory provision, Parliament did not seek to define public benefit or alter its pre-existing meaning. The alleged existence of a presumption of public benefit was posited on a statement made by Lord Wright, where the issue before him was the matter of whether the National Anti-Vivisection Society had been established for charitable purposes only.[96] Acknowledging the somewhat distinctive position of trusts falling under the first three *Pemsel* heads, he concluded that: 'The test of benefit to the community goes through the whole of Lord Macnaghten's classification, though as regards the first three heads, it may be prima facie assumed unless the contrary appears.'[97] Tellingly, he does not employ the term 'presumption', but alludes to an assumption that, in the absence of evidence to the contrary, a purpose in those categories is beneficial. Lord Simonds agreed that without that evidence, 'the court will easily conclude that it is a charitable purpose'.[98]

In the *ISC* case, however, the Upper Tribunal denied that a presumption of public benefit had ever existed, but acknowledged that there was a 'predisposition', which might be displaced by evidence, argument or a consideration of the purpose in the mind of the judge. Following such displacement, 'evidence would be needed to establish public benefit'.[99] The impact of this 'predisposition' was confined to the benefit aspect of public benefit and it did not serve to preclude the court from arriving at its own conclusions on the matter, 'not by way of assumption, but by way of decision'.[100] This interpretation is clearly arguable and can be made to fit any *ex post facto* analysis of the case law. Accordingly, the inherently educational benefits of a museum displaying antiques or works of art[101] may be offset by evidence that the objects displayed are of no inherent artistic or intrinsic value.[102] The publication of religious writings is readily deemed to advance religion unless the works under scrutiny might 'inculcate doctrines adverse to the very foundations of religion or be subversive of all morality'.[103] The issue is altogether different where the charitable character of a purpose is also in dispute and the judge may, understandably, require to be convinced of the quality of benefit. The classic illustration

94 Synge (n 28) 223.

95 S 4(2) provides that 'it is not to be presumed that a purpose of a particular description is for the public benefit'.

96 *National Anti-Vivisection Society v IRC* [1948] AC 31.

97 Ibid 41.

98 Ibid 65.

99 [2011] UKUT 421 (TCC) [67].

100 *AG v Charity Commission for England & Wales* [2012] UKUT 420 (TCC), [39].

101 *In Re Spence* [1938] Ch 96, 105, the display of arms and antiques in a public hall was deemed, 'an educational object' and, therefore, charitable in status.

102 *Re Pinion* [1965] Ch 85. As Wilberforce J commented at 95: 'Once a gift is clearly seen to be in the . . . educational sphere, it does not prevent the court from ascertaining, if necessary by evidence, whether the gift has any educational tendency.'

103 *Thornton v Howe* (1862) 31 Beav 14, 20.

is the trust for the training of spiritualist mediums,[104] cited by the Upper Tribunal in *ISC* as its primary example of a case where public benefit could not be assumed.[105]

In ISC, the strident disavowal of any technical common law presumption of public benefit generated surprise, at least in some quarters. In its amended analysis, the Commission acknowledged that 'it was widely considered that there was a presumption . . . although views differed as to the nature and effect'.[106] Those championing reform were convinced that the removal of every trace of a presumption would ensure the application of 'the same strong public benefit test across all four heads of charity'[107] and 'create a level playing field on which all charities will have to show that they are for the public benefit'.[108] For Luxton, however, s 4(2) had no effect on charitable status, 'because such status is not dependent on any presumption but has been established by law'.[109] On the basis that the section did not have the impact that was envisaged, the provision came to be viewed as no more than 'a political statement, which has affected practice and perception but hasn't really altered the law'[110] or as serving only 'to give expression to the extant practice of the Charity Commission, whereby all charities have been required to demonstrate public benefit in order to be registered as a charity'.[111] Sparked by the Commission's initial refusal to register the Preston Down Trust, the perceived difficulties to which religious charities were exposed were voiced loudly in the press[112] and in Parliamentary debate.[113] This led to the recommendation of the Public Administrative Select Committee that the 'removal of the presumption of public benefit . . . be repealed'.[114] The response of the government was unhelpfully trite, that 'it would not be possible to "restore" a presumption of public benefit that may never have existed'.[115] The continuing uncertainty prevalent in the discourse is far from satisfactory and the impact of the reform continues to be contested.[116]

The fourth head: presumption and fallacy

As previously indicated, for purposes within the new descriptions the prevailing view was that the test of public benefit would 'remain largely unaffected by the Act'.[117] Whatever the uncertainties facing charities in demonstrating public benefit for the first time, it was plainly understood that fourth head purposes would be untouched by s 4(2). Understandably, purposes under this *Pemsel* head were never viewed as benefiting from the

104 *Re Hummeltenberg* [1923] 1 Ch 237.

105 *ISC* (n 25) [67].

106 Charity Commission, *Analysis of the Law Relating to Public Benefit* (September 2013) [12].

107 *For the Public Benefit: A Consultation Document on Charity Law Reform* (NCVO 2001) [4.1.6].

108 Hilary Armstrong, HC Deb 26 June 2006, vol 448, col 24.

109 Luxton (n 70) 9.

110 *Final Report and Recommendations of NCVO's Independent Review of the Charity Act 2006* (NCVO May 2012) 26.

111 C Buckley, 'The Charities Act 2006: Consolidation or Reform?' (2008) 11(1) Charity Law and Practice Review 1, 41.

112 See, for example, 'Proposed Changes to Charitable Status' *The Times* (London, 27 May 2004) 27.

113 See HC 13 November 2012, vol 553, col 36. This was described by Nick Hurd (at col 59) as 'quite a rough day for the members of the Charity Commission'.

114 Public Administrative Select Committee, *The Role of the Charity Commission and 'Public Benefit': Post-legislative Scrutiny of the Charities Act 2006 Third Report of Session 2013–14* (June 2013) [93].

115 *Government Response to the Public Administration Select Committee's Third Report of 2013–14: The Role of the Charity Commission and 'Public Benefit': Post-legislative Scrutiny of the Charities Act 2006* (Cm 8700 September 2013) 14.

116 See, for example, R Meakin, 'Taking the Queen's Shilling: The Implications for Religious Freedom for Religions being Registered as Charities' (2017) 178 Law and Justice 57.

117 P Smith, 'Religious Charities and the Charities Act 2006' (2007) 9(3) Charity Law and Practice Review 57, 70.

presumption of public benefit.[118] The fourth head differed from the first three as 'unlike them it was not a charitable purpose – rather it merely indicated a quality (that of being beneficial to the community) that any purpose would need to possess in order to enter the ranks'.[119] As Lloyd LJ acknowledged, 'there is a difference, at least in the process of analysis, between identifying a purpose as charitable within that aspect of the classification, on the one hand, and doing so within one of the first three categories, on the other'.[120] For novel purposes, two separate issues fell to be addressed. The court would consider whether, in a general sense, the purpose carried any benefit or value. If so, the purpose was required to fall within the spirit and intendment of the Preamble to the 1601 Act.

Only novel fourth head purposes were made subject to a more exacting test of validity, with the need for a distinctive approach to fourth head purposes diminishing as the head became shaped by established groupings of cases.[121] If a purpose could be housed within an established sub-category, it could, in effect, be assumed to carry benefit in the same way as trusts falling under one of the first three heads. By way of illustration, the welfare of animals had become a well-established fourth head sub-category by the early twentieth century,[122] for the reason that 'discouragement of cruelty to lower animals . . . involved benefit to the community'.[123] Accordingly, in *Re Grove-Grady Hanworth* LJ was clear that the 'removal of cruelty to animals' (the primary concern of the testatrix) was an object that 'must clearly be accepted as a charity within the decided cases'.[124] This type of benefit is, however, intangible, unquantifiable and thoroughly anomalous. It is described in typically oblique fashion in the case law through the use of phrases such as the promotion of feelings of morality, the repression of brutality and the elevation of the human race.[125] Not every trust for the benefit of animals carries this form of benefit to the community. Indeed, Picarda admitted that: 'It is difficult to see how animal charities are within the spirit and intendment of the preamble.'[126] However, once the capability of benefit (in a legal sense) is recognised, that is, in the identification of analogous case law, the court is involved in exactly the same process of evaluation as would occur under any of the first three heads. Evidence of benefit thus ceases to be of concern and what follows is the taking on board of any countervailing evidence of lack of benefit. The outcome in *Re Grove-Grady* is well known. Lord Hanworth MR could find nothing beneficial to the animals or to mankind from a sanctuary 'free from the molestation of man, while all the fauna within it are to be free to molest and harry one another'.[127]

The desire for a strong, universal test of public benefit was built upon the fallacy that benefit under the fourth head could never be assumed, that such purposes were subject

118 As Baroness Scotland explained of the general law (HL Deb 20 January 2005, vol 668, col 885), 'there is a presumption that charities established for the relief of poverty, the advancement of education or the advancement of religion are for the public benefit. However, charities established for all other purposes do not benefit from this presumption.'

119 See Luxton, 'A Three-Part Invention' (n 76) 22.

120 *Helena Partnerships Ltd v HM Commissioners for Revenue and Customs* [2012] EWCA Civ 569, [65].

121 As J Hackney put it in 'Charities and Public Benefit' [2008] Law Quarterly Review 347, 348: 'these purposes became grouped together into classes, and though the language of the books does not always adequately reflect it, other 'heads' came into analytical existence'.

122 See *Re Wedgewood* [1915] 1 Ch 113.

123 Per Russell LJ in *Re Grove-Grady* [1929] 1 Ch 557, 588.

124 Ibid 572.

125 See *Re Wedgewood* (n 122).

126 Picarda (n 39) 109.

127 [1929] 1 Ch 557, 574.

to a higher degree of scrutiny and that more robust screening might also apply fittingly to purposes under the first three *Pemsel* heads. Accordingly, benefit is seemingly demonstrated by the weighing-up of benefit and detriment in a given case, as though the purpose at hand were a novel purpose, and without the assumption that a purpose falling squarely within a statutory description is a beneficial purpose. In its day-to-day assessment of public 'benefit', moreover, the Commission sets a high bar with an unapologetic focus on the means by which a purpose is pursued. The relevant question is, 'not whether the . . . [particular] purposes are *capable* of being for the public benefit, but whether there is sufficient evidence from which it can be concluded that they *will* result in such benefit' (original emphasis).[128] This is, most certainly, an exacting test for any newly established organisation that might struggle to furnish satisfactory evidence of impact. When challenged on its interpretation of the law in *Uturn*, the Commission justified its general approach by reference to s 2(1)(b) and the 'absolute requirement' now imposed that a purpose be for the public benefit.[129] In the context of community development, the Commission was forced to concede that its earlier, less rigorous statements to the contrary pre-dated the Act, did not reflect the current law and should be disregarded. It is far from clear that this strict and uncompromising approach to public benefit, which it applies across the board, is either justifiable or fair.

Public benefit guidance

Pursuant to its public benefit objective, the Commission is required to issue and, from time to time revise, its public benefit guidance.[130] This is guidance to which charity trustees must have regard when exercising any powers and duties to which that guidance is relevant.[131] Issued in 2013, the guidance in its current form entirely replaces the earlier 2008 version[132] and comprises three parts. Whereas parts two and three deal with running a charity and reporting on a charity's work respectively,[133] only the first part (PB1) directly concerns the 'public benefit requirement' as defined in the 2011 Act.

To satisfy the 'benefit aspect' of public benefit, the guidance makes clear that 'the purpose must be beneficial' and 'any detriment or harm that results from the purpose must not outweigh the benefit'.[134] Beyond this broad prescription, the Commission guidance is brief, vague and pitched at a level of generality that offers little insight into the Commission's expectations, method of working or rigorous quality control. Its choice of illustrations is mundane and unchallenging. It clarifies that beneficial means 'beneficial in a way that is identifiable', 'capable of being proved by evidence where necessary' and 'not based on personal views'.[135] The reference to 'where necessary' is immediately eye-catching. The Commission admits that 'In some cases the purpose is so clearly beneficial that there is little need for trustees to provide evidence to prove this',[136] chiming clearly

128 *Uturn* (n 56) [40].

129 Ibid [41].

130 S 17.

131 S 17(5).

132 *Charity Commission, Charities and Public Benefit* (January 2008).

133 *Public Benefit: Running a Charity (PB2)*; *Public Benefit: Reporting (PB3)*.

134 *Charity Commission, Public Benefit: The Public Benefit Requirement (PB1)* (September 2013) 5. Based on evidence and not personal views, the Commission will: 'take detriment or harm into account where it is reasonable to expect that it will result from the individual organisation's purpose' (8).

135 Ibid 7.

136 Ibid 8, the Commission noted that, 'where the benefit of a purpose is obvious and commonly recognised, there is an even greater need for evidence of detriment or harm to be clear and substantial, if it is to outweigh that benefit'.

with the view expressed by Vaisey J that, 'there are many cases . . . where the purpose is so obviously beneficial to the community that to ask for evidence would really be quite absurd'.[137] It is unhelpful, nonetheless, that the Commission provides no more than a solitary example of the most urgent and compelling kind, 'the provision of emergency aid in the context of a natural disaster'. Given that a great many trustees may feel that their purposes are clearly beneficial without the need for probative evidence, it is regrettable that no borderline illustrations are offered to assist. This part of the guidance also fails to convey that the focus of the Commission falls not only on the purpose, but the manner in which the purpose is pursued, leading it to doubt, more often than it might, that benefit will result from the pursuit of the organisation's purposes.

Obvious and other benefits

The requirement for formal proof of benefit was contested in the *HDT* case. HDT had been denied registration by the Commission on the basis that its objects were too vague and its purposes too political. On appeal, the Commission argued that 'the onus of proof fell on HDT to satisfy the Tribunal that its purposes are directed to benefitting the public in a way that is recognised as charitable'.[138] Accordingly, HDT argued that the eradication of human rights abuses was so obviously beneficial that it did not demand formal proof. Support for this view was found in the Commission's own RR12 guidance which stated that the 'concept of human rights is virtually unanimously endorsed by the countries of the world'[139] and that 'there is an obvious benefit in promoting human rights'.[140] Although the Tribunal did not respond specifically to that argument, it was satisfied by witness evidence presented by HDT which was couched in the language of particular benefits to individuals whose human rights were promoted and protected and the wider benefits to the community at large from having such rights interpreted, clarified and enforced. No evidence of detriment was presented. The Tribunal concluded readily that the 'criminalisation of relevant conduct represents a serious breach of human rights norms' and that there must be 'a public benefit in seeking to interpret, clarify and protect superior constitutional rights'.[141]

To clarify cases where the Commission may ask for evidence, the guidance provides straightforward examples, such as to determine the artistic merit of an art collection or the healing benefits of a therapy. These are uncontroversial and unsurprising and are clearly the type of claims that will necessitate expert evidence to establish their merits. Such evidence was adduced by the Soteria Network as to the efficacy of its alternative healing method.[142] Organisations that fail to do so will simply fall at the first hurdle. In the application for registration by Hitchin Bridge Club,[143] the Club sought to be considered under the advancement of amateur sport, which embraces a sport or game which promotes 'health by involving physical or mental skill or exertion'.[144] To fall within the description, it was necessary to demonstrate associated health benefits of the game and, in reaching its decision, the Commission took account of a body of research

137 *Re Shaw's WT* [1952] Ch 163, 169.

138 CA/2013/0013 (9 July 2014) (FTT (Charity)) at [102].

139 Charity Commission, *RR12: The Promotion of Human Rights* (January 2005), [11].

140 Ibid [12].

141 *HDT* (n 44) [109].

142 See the Decision of the Charity Commission on Soteria Network (1 December 2012) <www.gov.uk/government/publications/soteria-network>.

143 Hitchin Bridge Club, 1 March 2011 <www.gov.uk/government/publications/hitchin-bridge-club>.

144 S 3(2)(d).

evidence which identified those benefits. The 'benefit' element was established just as soon as it was deemed to fall within the confines of the description. In its separate assessment of public benefit, the Commission had little more to add than a series of generic observations, such as the undoubted benefit of, 'regular mental stimulation' from participation in the game.[145] The *Cambridgeshire Target Shooting Association*, however, did not fare so well.[146] The appellant introduced witness evidence from an experienced target shooter, a retired clinical psychologist, a specialist sports physiotherapist, a reader in human neuropsychology and a consultant physical and chemical pathologist. The claim was undone by the absence of published, discrete scientific research. The Tribunal warned that neither it nor the Commission 'should lightly reach findings, which result in the advantages of charitable status being conferred in respect of activities for which the science base is less than robust'.[147] Presumably, if the Association had demonstrated its rightful place within the description, its inherent benefit would have been established without more ado.

Evidential difficulties

The final advice in this part of the Commission's guidance deals with the difficulty of measuring benefit and maintains that it should 'always be possible to identify and describe how a charity's purpose is beneficial, whether or not that can be quantified or measured'.[148] The example provided is 'developing a person's artistic taste by viewing works of art'. There would be few who would contest the benefit of such a purpose were the works of art of sufficient merit. The greater concern, which arises where some formal evidence of benefit is demanded, focuses upon the quality of evidence required. In its separate *Analysis of the Law Relating to Public Benefit*,[149] the Commission acknowledged that: 'The courts have not been uniform in their use of language to describe the test for deciding whether a purpose is for the public benefit.'[150] It added that, 'the quality of the evidence which is required to satisfy this test depends on the circumstances of the particular case'. It follows that detailed guidance is difficult to compile and the outcome of an application to the Commission difficult to predict. In *Full Fact*, for example, the Commission could be satisfied by nothing less than evidence of external validation and concrete reassurance as to the future operation of the organisation. Its eventual registration took five years to secure and required both the adoption of, 'constitutional provisions for an independent reviewer to periodically assess its work'[151] and the provision of supporting references from the UK Statistics Authority, Royal Statistical Society and The Migration Observatory. In this case, the evidence that followed was both robust and compelling.

More interesting, however, are those cases where the Tribunal has reached a different conclusion to the Commission either on the basis of the evidence presented to it or, on occasion, the absence of evidence negating benefit. These decisions serve best to illustrate the fundamental difference in outlook between the two bodies. In striking contrast to the Commission, the Tribunal elects for a broad-brush approach to the

145 Hitchin Bridge Club (n 143) [28].

146 *Cambridgeshire Target Shooting Association v Charity Commission for England and Wales* CA/2015/0002, 23 November 2015 (FTT (Charity)).

147 Ibid [61].

148 *Public Benefit: The Public Benefit Requirement (PB1)* (September 2013) 7.

149 Charity Commission (n 106).

150 Ibid [50].

151 CA/2011/0001, 26 July 2011 (FTT (Charity)) [31].

demonstration of benefit, which allows for intuitive responses and much greater flexibility. In *Uturn*, for example, the First-tier Tribunal acknowledged that the Commission may have been influenced by the absence of evidence of demonstrable benefit. Some five months on from the Commission decision, the Tribunal seemed prepared to accept that satisfactory evidence of benefit could have come first-hand from 'individuals actually participating in (or benefitting from) the activities'[152] and described it as 'regrettable' that that growing evidence of beneficial impact was not supported by witness evidence. Admittedly, neither the Commission nor the Tribunal found that Uturn was established for exclusively charitable purposes, and so the question of public benefit was considered by the Tribunal for the sake of completeness. It is difficult to imagine, nonetheless, that such witness evidence would have satisfied the Commission had it been presented at the time. The Commission's evident preference for robust external validation would seem to go much beyond the subjective or self-serving statements of those participating in the activities.

In *HOPRT*, the Commission and the Tribunal reached contrasting conclusions as to the purposes for which HOPRT was established and on the matter of public benefit. Once the Tribunal was satisfied that HOPRT's particular research-orientated purposes were charitable within the parameters of education, little was deemed to be required by way of formal proof to evidence its benefit. It certainly helped that only minimal evidence was presented to the contrary. In terms of its public benefit assessment, the Tribunal formulated the key question to be whether the research conducted by HOPRT was 'in its broadest sense, a *good thing*' (original emphasis).[153] Unchallenged witness evidence was presented to this end, describing benefits in the broadest of terms, such as the giving of comfort to patients and making a mother's death easier to bear. Acknowledging that chances of reanimation were remote, there was little to negate the benefits of the research conducted by HOPRT. Ethical uncertainties had not been an issue in the appeal and no evidence on the matter was presented. In addition, there was no evidence of pressure or false hope for those in receipt of services. As cryopreservation was lawful, unregulated by Parliament and in demand, the Tribunal held that the benefit part of the public benefit requirement was clearly satisfied.

Finally, in the *IPRT* case,[154] the First-tier Tribunal swept aside the concerns of the Commission that there was no evidential basis from which the public benefit of a Leveson-compliant system of press regulation might be discerned.[155] Instead, it again utilised a broad-brush approach without regard for remoteness or future uncertainties. The Tribunal accepted evidence in the form of a written report outlining the public benefit. The founding director of the IMPRESS project[156] lacked the impartiality to be deemed an expert witness but gave uncontested evidence as to the public benefit of the project and the Tribunal happily endorsed his views. It noted that recommendations had been made in the wake of a thorough public enquiry and that those recommendations

152 CA/2011/0006, 27 February 2012 (FTT (Charity) [42]. The Tribunal was receptive to the fact that a further five street associations had been set up by the date of the hearing and heard descriptions of the benefit that flowed from them, namely new links formed in participating communities and the encouragement of volunteering.

153 *HOPRT* (n 78) [120].

154 CA/2014/0022, 15 June 2015 (FTT (Charity)).

155 See the Decision of the Charity Commission on the Independent Press Regulation Trust (29 October 2014) [25] <www.gov.uk/government/publications/independent-press-regulation-trust-iprt>.

156 Another organisation seeking to establish a similar press regulator.

retained the support of all parties in Parliament. Apparently influenced by matters of public policy, the Tribunal declared that:

> . . . if such a regulator cannot be established by the Government for constitutional reasons and ought not to be established by the industry itself for reasons of propriety and public confidence, then the charity sector is uniquely placed to be able to offer both the mechanism and the means by which a benefit to the community as a whole can be achieved.[157]

Such a conclusion was firmly outside the Commission's view of its remit and range of responses.

Conclusion

This article has identified various problems inherent in redefining the concept of a charitable purpose, specifically from the perspective of the old fourth head. As demonstrated, some of those problems are conceptual, deriving from the lack of established meaning in the statutory language and the open-endedness of the descriptions themselves. Others stem from the elusive concept of public benefit and the strict manner in which it is now interpreted and applied. Understandably, the Commission must approach its task in a robust fashion. It is aware of its role as a gatekeeper, securing the charity brand and guarding officiously the fiscal privileges and reputational advantages attendant to charitable status. In its insistence upon compelling formal proof of impact and invariable focus on activities, there is no doubt that it sets a challengingly high standard that otherwise worthy applicants have struggled to meet. There lies the major obstacle to the fair and inclusive operation of the test for charitable status.

By contrast, the First-tier Tribunal operates with freedom and flexibility. In considering afresh the claim for eligibility to charity registration, the Tribunal does not set out to identify or clarify the errors of the Commission and is able to take into account evidence which was not available to the Commission.[158] Although not a body generating precedent, the Tribunal's decisions may undoubtedly carry some weight.[159] Such was the case with *HDT*, viewed conservatively by some as a 'logical application of the existing law in this field',[160] but for others as 'most significant', making it 'increasingly difficult to sustain the political purposes exclusion where . . . human rights are at issue'.[161] Unfortunately, many of the appeals to the Tribunal have failed to stimulate any meaningful comment. With so few cases now reaching a court of record, serious doubt is cast as to the scope for charity law to develop organically and systematically.

For the applicant, the path to charitable status is daunting and unpredictable. There is an unhelpful absence of developing precedent and little can be gleaned from the Register of Charities as it 'does not exist as a precedent bank'.[162] It seems to make no difference that an applicant can point to examples of charities registered with purposes that are the similar or analogous to its own. It is clear too that the Commission's subject-specific non-

157 *IPRT* (n 72) [40].

158 Charities Act 2011, s 319(4).

159 In the words of A McKenna, these cases offer 'a fascinating snapshot of what is really going on': 'How Does Charity Law Develop in the Age of the Tribunal' (2018) 20 Charity Law and Practice Review 25, 32.

160 P Kirkpatrick, 'Human Dignity Trust v Charity Commission for England and Wales' (2015) 2 Private Client Business 73, 73.

161 H Biehler, 'The Political Purposes Exception – Is There a Future for a Doctrine Built on Foundations of Sand?' (2015) 29(3) Trust Law International 97, 102.

162 *Crocels Community Media Group v Charity Commission for England and Wales* CA/2015/0009, 18 July 2016 (FTT (Charity) [23].

statutory guidance cannot be relied upon. Although its work on the Register has been bold and confident,[163] much of the guidance produced is of its time and does not help to impart 'particular meaning' in the statutory concepts. The authoritative steer is to be found in the, 'previous (binding) decisions of the court rather than . . . [the] guidance issued by a regulator'.[164] For the disappointed applicant, the First-tier Tribunal may now offer a viable avenue of appeal with reasonable prospects of success. Although manifestly cheaper than undertaking High Court litigation, the price of access to justice in the First-tier Tribunal may nevertheless be off-putting in terms of the time and effort required to engage with the appeal process and, as will often prove necessary, the associated cost of instructing counsel.[165]

163 See Charity Commission, *The Review of the Register of Charities (RR1)* (1 October 2001); Charity Commission, RR1A: *Recognising New Charitable Purposes* (1 October 2001).

164 *HDT* (n 44) [42].

165 No better illustrated than in *Full Fact* (n 52) where the registration process took a full five years and cost the organisation thousands of pounds in legal fees.

NILQ 70(4): 467–477

The Stephen Livingstone Lecture: 'The problems with human rights' – given on 21 November 2019

Brice Dickson

Queen's University Belfast

Abstract

This is the text of the 2019 Stephen Livingstone Lecture delivered at Queen's University Belfast on 21 November 2019. It explores three types of problems which frequently arise when advocates of human rights try to convert a claim into a human right. These are philosophical problems (people differ greatly in how they conceptualise human rights), legal problems (it is an accident of legal history that human rights became a term of art in international law before it did so in national law, meaning that even today some human rights activists maintain the view that only states can be accused of violating human rights) and practical, or implementational, problems (where there difficulties with the remedies made available to victims of human rights abuses).

Keywords: human rights; conceptions of rights; Bill of Rights; non-state actors

Introduction

Thanks very much for that kind introduction. I'd like to begin by reiterating very sincerely that those of us in the Law School who knew Stephen still miss him very much. His intellect and his good humour are unforgettable, and it is therefore all the more a privilege and a pleasure to be again giving the lecture established in his name. I only hope I can do his memory justice.

The title of my talk is 'The problems with human rights'. It was partly chosen as a provocation, an attempt to lure an audience. But I want to make it clear at the outset, lest there should be any misunderstanding, that I am not here to argue that human rights are themselves a problem. Far from it, for I remain a strong supporter of human rights. Instead, I want to address the question of why it is often so difficult for those who are advocates for human rights to ensure that human rights are appropriately protected *by law*. I want to suggest that there are basically three types of problem which hinder the achievement of that goal. I'll cover each of them in detail, but let me start by listing what they are (and please bear in mind that the categories overlap a little).

Firstly, there are *philosophical problems*. These arise from the fact that, when asked to really think about it, people tend to differ over what they think human rights are. Their conceptualisation of human rights varies greatly. Putting this another way, people disagree over what human rights *are for*.

Secondly, even if we could all agree on how to conceptualise human rights (and I do not think that is at all likely), there are still problems over how to express human rights *in law* – i.e. *legal problems*. Which human rights are already well recognised in law? What are the limits to those rights? And who should be held accountable for breaches of those rights?

Thirdly, even if the law precisely defines a human right, there are often problems associated with how to protect that right: what mechanisms are available for identifying whether it has been violated, and what remedies should be provided to victims of the violation and to prevent similar violations in the future? I am calling these *practical (or implementational) problems*.

So, having roughly mapped out my territory, let me spend the next 35 minutes or so reflecting in more detail on the various problems.

Philosophical problems

There is a huge literature on this topic and innumerable accounts have been presented as to what really lies at the base of the concept. A Professor of Law and Anthropology at the University of Sussex, Marie-Bénédicte Dembour, has written a great book entitled *Who Believes in Human Rights?*,[1] where she examines four of the myriad ways in which human rights can be conceptualised. She first discusses the utilitarian approach, according to which a human rights claim should be supported if doing so adds to the sum total of human happiness in society. She then explains the Marxist approach, which supports the protection of individual rights only if doing so ensures that every member of society gives according to his or her abilities and receives according to his or her needs. The cultural–relativist approach maintains that in different societies at different times particular human rights should be recognised, such as the right to engage in certain rituals, gather in certain places or commemorate certain deeds. Finally, Dembour expounds the feminist approach to human rights. This seeks to redress actual and perceived imbalances in the way human rights are enjoyed by men but not by women, as well as claiming protection for rights that only women could have, such as the right to terminate a pregnancy.

There are numerous other ways of conceptualising human rights. Some people are convinced that they derive from morality, whether the content of that morality is determined divinely (which is the so-called natural law approach, still supported by some judges and legal academics in the Republic of Ireland, by the way) or whether the morality derives from another core value such as 'personhood', as suggested by the American philosopher James Griffin,[2] or 'dignity', something which my colleague Christopher McCrudden has written about.[3]

Today's so-called 'critical legal scholars' tend to conceptualise human rights as a kind of sop, granted to individuals to make up for the negative aspects of capitalism, colonialism and globalisation. The Kenyan-born scholar Makau Mutua talks sarcastically about human rights being something which the liberal West has generously 'gifted' to the developing world. He castigates human rights activists for portraying themselves as 'saviours', riding to the rescue of victims who have been deprived of their agency by

1 Cambridge University Press 2010.

2 *On Human Rights* (Oxford University Press 2009).

3 E.g. 'Human Dignity and Judicial Interpretation of Human Rights' (2008) 19 European Journal of International Law 655–724. See too C McCrudden (ed), *Understanding Human Dignity* (British Academy 2013).

'savages'.[4] Twenty years ago the Harvard-based scholar David Kennedy suggested that things were already so awry that the human rights movement had itself become part of the reason why there were still so many serious human rights violations in the world.[5]

The most recent critiques of human rights focus on the failure of the concept to deal with the worldwide problems of poverty, inequality and climate change. These frustrations are evident in book titles such as Eric Posner's *The Twilight of Human Rights Law* (Oxford University Press 2014), Stephen Hopgood's *The Endtimes of Human Rights* (Cornell University Press 2015), and Samuel Moyn's *Not Enough: Human Rights in an Unequal World* (Belknap Press 2018).

The common thread running through many of these critiques is that protecting individual rights, while it may do a lot to protect the interests of liberal elites, is not successful at tackling deep-rooted structural flaws in the international world order. In the year 2000, in part acknowledgment of that reality, the UN adopted a set of eight Millennium Goals, to be reached by 2015.[6] They included the eradication of extreme poverty and hunger, the delivery of universal primary education, the promotion of gender equality and the reduction of child mortality. Needless to say, the goals were not reached within the timescale envisaged, so in 2015 a further set of 17 Sustainable Development Goals was adopted, targeted for achievement by 2030.[7]

It is too early to say whether the focus on setting goals and changing policies has had a greater impact on systemic human rights violations than the international human rights movement has had, but that may well turn out to be the case. Of course, the one approach does not preclude the other. A human rights approach can still have some kind of effect on the systemic problems in question. My colleague Kathryn McNeilly, in her challenging book entitled *Human Rights and Radical Social Transformation*,[8] makes a plausible case for the view that a human rights approach to fundamental problems *can* facilitate radical social change.[9]

I've briefly surveyed different ways in which human rights can be conceptualised in order to demonstrate how problematic it can be for human rights advocates to gain

4 'Savages, Victims, and Saviors: The Metaphor of Human Rights' (2001) 42(1) Harvard International Law Journal 201–245. See too T Evans, 'International Human Rights Law as Power/Knowledge' (2005) 27 Human Rights Quarterly 1046–68; C McCrudden, 'Human Rights Histories' (2015) 35 Oxford Journal of Legal Studies 179–212.

5 'The International Human Rights Movement: Part of the Problem? (2001) 3 European Human Rights Law Review 245; reprinted at (2002) 14 Harvard Human Rights Journal 101–126. He updated this in a later book chapter: 'The International Human Rights Regime: Still Part of the Problem' in O W Pedersen (ed), *Examining Critical Perspectives on Human Rights* (Cambridge University Press 2013).

6 There were eight of them: (1) to eradicate extreme poverty and hunger; (2) to achieve universal primary education; (3) to promote gender equality and empower women; (4) to reduce child mortality; (5) to improve maternal health; (6) to combat HIV/AIDS, malaria and other diseases; (7) to ensure environmental sustainability; and (8) to develop a global partnership for development.

7 There are 17 of these, the targeted completion date being 2030. They include clean water and sanitation, affordable and clean energy, decent work and economic growth, responsible consumption and production. See <www.un.org/sustainabledevelopment/sustainable-development-goals>.

8 Routledge 2017. The author's approach is premised on the idea that human rights can be reconceptualised 'as inherently futural and capable of sustaining a critical relation to power and alterity in radical politics'.

9 A not dissimilar approach has been advocated in several articles by Paul O'Connell from SOAS in London, although his manifesto is more avowedly Marxist in tone: see 'On the Human Rights Question' (2018) Human Rights Quarterly 962–988; 'Brave New World. Human Rights in the Era of Globalization' in M Baderin and M Ssenyonjo (eds), *International Human Rights Law: Six Decades after the UDHR and Beyond* (Taylor & Francis 2016) ch 10; 'Human Rights: Contesting the Displacement Thesis' (2018) 69 NILQ 19–35; 'The Death of Socio-Economic Rights' (2011) 74 Modern Law Review 532–54.

support for human rights claims, especially novel ones. Not everyone shares the same idea of what a human right is. I often wonder whether unionists and nationalists in Northern Ireland have the same concept in their heads. My next-door neighbour in the Law School, Louise Mallinder, recently published a piece which argues that unionists have adopted strategies towards the application of human rights law which are aimed at promoting their own political interests.[10] I dare say similar evidence could be adduced to show that nationalists have done the same. It is interesting at the moment to see how unionists and nationalists are reacting differently to suggestions that some kind of line should be drawn under all investigations, prosecutions, inquiries and inquests relating to the troubles in Northern Ireland. There seems to be increasing support for those suggestions in the very highest corridors of power and spheres of influence. There are, indeed, human rights arguments for and against drawing such a line. Back in 1996, for example, the Constitutional Court of South Africa ruled that even access to justice has to take a back seat if what is in the front seat is the reconciliation of a country's people.[11]

But, to end this brief examination of philosophical problems relating to human rights, let me run by you some of the recent candidates proposed for human rights status so that you can reflect in your own time on whether they do deserve to be recognised as human rights: the right of childless couples to receive free IVF treatment; the right of sex offenders to receive chemical castration treatment (because they genuinely want to have their sex drive reduced);[12] the right of parents to have their baby sons circumcised; the right of parents to have their daughters' genitals cut; the right to die with dignity; the right to have free access to the internet; the right of prisoners to vote in elections; the right of 16- and 17-year-olds to vote in elections; the right 'to be forgotten' (so that internet search engines cannot continue to spew out information about your past, even if it is true);[13] the right to a uniform basic income; the right to migrate to whichever country you want to go; the right not to be discriminated against merely on the basis of where you live in a country; the right to a higher life expectancy than your parents had; and, last but by no means least, the right to clean air. That is quite a mixed bag, and I would suggest that where you stand on whether any of those claims is a genuine human right or not will depend very much on your understanding of what the purpose of a human right actually is.

Legal problems

Let me turn now to some legal problems facing human rights advocates. The first of these, paradoxical though it might seem, is that 'human rights' became a term of art in international law before it did so in national law, even though it was in the national law of France that the concept first saw the light of day. The legal document which broke new ground was the Declaration of the Rights of Man and the Citizen endorsed by the French National Assembly on 26 August 1789. It continues to be part of France's Constitution today and still makes for inspiring reading. It highlights just four specific rights – the right to liberty, to property, to security and, most surprisingly perhaps, to resist oppression. It defines liberty as the permission to do anything you want to do so

10 'Metaconflict and International Human Rights Law in Dealing with Northern Ireland's Past' (2019) 8 Cambridge International Law Journal 5–38.

11 *AZAPO v President of the Republic of South Africa* 1996 (4) SALR 671 (CC).

12 This has been an issue in the Czech Republic.

13 See the decision of the Court of Justice of the European Union (CJEU) in *Google Spain SL v Agencia Española de Protección de Datos and González* (Case C-131/12; 13 May 2014). For commentary on two recent CJEU decisions expanding upon the 2014 decision, see <https://privacylawblog.fieldfisher.com/2019/the-right-to-be-forgotten-and-the-eu-court-of-justice-round-2>.

long as it does not harm anyone else. Two years later a more specific Bill of Rights was adopted by the young USA, and that document too is very much alive and well, although heavily amended. The Americans call those rights constitutional rights, not human rights. And for over 70 years, let us remember, America lived with its Bill of Rights while at the same time tolerating the practice of slavery. English common law, if we are honest about it, has never fully embraced the idea of human rights. From the seventeenth century onwards it recognised certain 'liberties' and 'freedoms', but it was not until the second half of the twentieth century that UK judges began to call these 'rights', and even today it is rare for a judge to categorise a common law right as a human right or a constitutional right; the furthest the judge might go is to label it a fundamental right.

What changed things dramatically for the concept of human rights was the world's reaction to the horrors of Nazism in the 1930s and 1940s. The Universal Declaration of Human Rights, issued on 10 December 1948, marked a seismic shift. In just 30 articles, and fewer than 1500 words, the United Nations set out a list of human rights 'to the end that every individual and every organ of society . . . shall strive . . . to promote respect for these rights . . . and by progressive measures . . . secure their universal and effective recognition and observance'.[14]

The Universal Declaration has become the most translated document in history, now available in more than 500 languages. It asserts that human rights are values which need to be guaranteed by states to all individuals living in those states in order to preserve a minimum level of human decency. This was a huge development because until then international law required states to be beholden only to other states, not to any individuals. But, of course, the Declaration was only an *inspiring statement*, not a *binding agreement*, and it took a further 28 years before its provisions became obligatory for those states which had ratified the two subsequent UN treaties, or Covenants, one on Civil and Political Rights and the other on Economic, Social and Cultural Rights. Today those two Covenants have been ratified by 173[15] and 170[16] states respectively, though the fact that China has not ratified the Civil and Political Covenant and the USA has not ratified the Economic, Social and Cultural Covenant tells you a lot about how states differ in the way they prioritise different rights. Moreover, in states which have ratified both Covenants, civil and political rights are generally much better protected than socio-economic rights.

A further point is that, very commendable though they are, the two UN Covenants reinforced the idea that human rights are rights which can be claimed only against states and not against any other entity or individual. That was not an inevitable deduction from the Declaration, which is specifically addressed not just to all peoples and nations but also, as I have said, to 'every individual and every organ of society'. The Covenants could have, but did not, embed the idea that human rights are *values* which need to permeate all levels of society and be legally vindicated even in disputes between private entities. The internationalisation of human rights law has therefore given the impression that only states can violate human rights, not anyone else – not non-state armed groups, not multinational companies, not trade unions, not the press, not you, I or any other individual.

This unintended consequence of the two Covenants has been exacerbated because the UN has developed seven additional 'core' human rights treaties which adopt the same

14 These are the last words of the Preamble to the Universal Declaration of Human Rights.

15 Not e.g. Malaysia, Myanmar, Saudi Arabia or the UAE. China has signed but not ratified.

16 Not e.g. Malaysia, Myanmar, Saudi Arabia, the UAE, Botswana or Mozambique. The USA has signed but not ratified.

approach. In chronological order they deal with racial discrimination (1965), discrimination against women (1979), torture and ill-treatment (1984), children's rights (1989), the rights of migrant workers and their families (1990), the right to be protected against enforced disappearance (2006) and the rights of persons with disabilities (2006).

And it is not only the UN which has sponsored human rights treaties. Amongst the other intergovernmental bodies to have done so are the Council of Europe and the EU. The treaty which is most relevant to the islands of Britain and Ireland, of course, is the European Convention on Human Rights (the ECHR).

So, with all of these international human rights treaties in place, it is easier than ever to identify the human rights that can be claimed against states under international law. And that is a very good thing. But international law is largely silent on the responsibility of entities other than states to protect human rights.

Some critical legal scholars are not happy with that conclusion because they maintain that the international legal system is itself the product of overt and covert power structures, as well as of cynical political manipulation. Oona Hathaway at Yale University has demonstrated that many governments ratify human rights treaties simply to improve their image in the world and without any intention of properly implementing the treaties.[17] My answer to such critics is that, whatever its faults, the international human rights system has still achieved a huge amount since the 1960s, and without it the victims of appalling mistreatment would be even more numerous today than they already are.

Before leaving the topic of international human rights law, it is worth mentioning the huge problem of when does a state's non-compliance with its human rights obligations become so grave that it is acceptable under international law for other states to use force against that state to protect the people living there from suffering further violations? Let us not forget that in the 1970s the world refused to use force to stop either Idi Amin from slaughtering some 500,000 people in Uganda or the Khmer Rouge from massacring more than a million-and-a-half people in Cambodia. The UN did not cover itself in glory prior to the killing of about 800,000 people in Rwanda in 1994, and UN troops shamefully stood by while 8000 Muslim Bosniaks were murdered by Serbian forces at Srebrenica in 1995. This year's review into the UN's inaction prior to the ethnic cleansing of the Rohingya people in Myanmar in 2017 also makes for depressing reading.[18]

To achieve reform in this area we clearly need changes to the UN's Charter and/or to its veto system. Alternatively, a brand new multilateral treaty needs to be negotiated on the protection of people against grave and systemic human rights violations by their own governments. In 2005 the UN *did* unanimously commit to a 'Responsibility to Protect' doctrine,[19] which was aimed at preventing genocide, ethnic cleansing, war crimes and crimes against humanity, but, alas, this commitment has not yet made the world a noticeably safer place. The people of Turkey, Syria and Yemen can certainly testify to that – those that are still alive that is.

Turning now to legal problems with human rights closer to home, let us consider the uncomfortable truth that, sometimes, human rights conflict with each other or need to be

17 'Do Human Rights Treaties Make a Difference?' (2002) 111 Yale Law Journal 1935–2042, online (revised 2015) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=311359>.

18 Gert Rosenthal, *A Brief and Independent Inquiry into the Involvement of the United Nations in Myanmar from 2010 to 2018*, 29 May 2019 <www.un.org/sg/sites/www.un.org.sg/files/atoms/files/Myanmar%20Report%20-%20May%202019.pdf>.

19 The 2005 World Summit Outcome Document, paras 138 and 139 <www.unsystem.org/content/2005-world-summit-outcome-document-16-september-2005>.

suppressed in order to allow other values to flourish. In the case of *Child S* in 2004, the House of Lords held that an eight-year-old boy's right to anonymity had to take second place to the right of the public to know that the boy's mother was being tried for killing his older brother.[20] During election periods, as we know only too well at present, what can or cannot be broadcast on radio or television has to be restricted in order to preserve the fairness of the electoral process,[21] and, of course, political advertising on those media is banned at all times (though not on the internet).[22]

I think our law has adopted a sensible approach when considering such limitations to rights. It requires the limitation to be in pursuit of a legitimate aim, to be necessary in a democratic society and to be proportionate, although reasonable people can reasonably disagree over at least one of those requirements, namely the test of proportionality. A good example is Belfast's 'gay cake' case between Gareth Lee and Ashers Bakery, where the right of someone not to be discriminated against on the basis of his sexual orientation or political opinion needed to be weighed against the right of a business not to be forced to express a political message it did not agree with. Where to strike that balance divided human rights activists up and down the country, but at the end of the day the five Supreme Court Justices who dealt with the dispute, including the two most liberal Justices – Lady Hale and Lord Kerr – came down on the side of the bakery.[23] Mr Lee, however, has lodged an application with the European Court of Human Rights, and we await with interest what view that court will take, probably within the next year.

We are now so inured to the way the European Court operates that we forget it was a total novelty when it was established 60 years ago this year. Until then no state could be held to account in an international court by an individual, only by another state. The impact of the European Court's case law has been huge, especially in the UK, where the ECHR was effectively made part of UK domestic law when the Human Rights Act 1998 came fully into force in October 2000. But, again, a possible downside of these welcome developments is that they have skewed still further the concept of human rights as one which pertains only to bad things done by states. In the European Court and under the Human Rights Act it is always the state, or one of the state's public authorities, which is the party trying to justify the alleged breach of rights. The extension of international law to individuals' complaints has not been matched by an extension to the range of potential respondents to those complaints.

It is easy to see how we've got to this position. Human rights law is a relatively new kid on the block, and it developed at the international level because there was a huge gap in the mechanisms available for holding states to account for mistreating their own people. Mistreatment of people by other entities was dealt with by national law under the headings of criminal law and civil law. Those laws were not portrayed as protecting human rights, but, in reality, that is what they were doing.

So, even though national law does not explicitly label murder, assault or damage to property as breaches of human rights *law*, they are to my mind still clear breaches of human rights *values*. It is therefore hypocritical, or at the very least an example of highly selective reasoning, for someone to call what a representative of the state does when he or she kills or injures someone unjustifiably (i.e. not in self-defence or in defence of

20 *In re S (A Child)* [2004] UKHL 47, [2005] 1 AC 593.

21 *R (ProLife Alliance) v BBC* [2003] UKHL 23, [2004] 1 AC 185.

22 *R (Animal Defenders International) v Secretary of State for Culture, Media and Sport* [2008] UKHL 15, [2008] 1 AC 1312.

23 *Lee v Ashers Baking Company Ltd* [2018] UKSC 25, [2019] 1 All ER 1.

others) a violation of human rights but to refuse to use that expression when talking about a killing or injury perpetrated by a private individual or organisation. Human rights NGOs were slow to pick up on this point until comparatively recently, but it is now mainstream thinking that non-state actors commit human rights violations too. There are already several books analysing such violations.[24]

In two fields, in particular, human rights are now applied in private contexts where they used not to be. One is the field of domestic violence. A senior lecturer in this Law School, Ronagh McQuigg, has written extensively on how a human rights approach to domestic violence has the potential to greatly reduce its incidence.[25] International treaty law now recognises that what a victim of domestic violence suffers is a violation of his or her rights and the state in question is therefore under a duty to try to prevent such violence and thoroughly investigate it when it occurs.

The other new field is that of businesses and human rights.[26] In view of the significant power of many of today's multinational corporations, some of them much wealthier than sovereign states, it would be perverse not to require them too to protect human rights. The shocking behaviour of the Union Carbide Corporation, which led to the loss of some 2000 lives in the Indian city of Bhopal in 1985, was an important turning point in this context, as was the UN's endorsement of the Ruggie Principles in 2011.[27] But, as we saw with the collapse of the Rana Plaza building in Bangladesh in 2013, some commercial enterprises continue to pay little heed even to the right to life of their employees. More than 1100 people died in that calamity.

A remaining challenge in the field of non-state actors and human rights is to find a way of ensuring that armed groups which violate human rights can somehow be made accountable for their violations. The International Criminal Court is beginning to make some progress on that front, but only slowly. A decision of the UK's Supreme Court last week suggested that the UN's Convention Against Torture might be enforceable against a member of a non-state armed group if it was effectively acting as a governing authority in a particular community at the time.[28] Even in the absence of legal proceedings, however, it is already perfectly proper to describe what non-state armed groups do as violations of human rights. There may well be a human right to resist oppression, but limits must still be placed on the kinds of action that can be justified under that heading.

Practical problems

I now turn, finally, to the third type of problem which besets human rights – practical problems. I want to highlight three of these.

24 See the books by Nadarajah Pushparajah, *Human Rights Obligations of Armed Non-State Actors in Non-International Armed Conflicts* (Wolf Legal Publishers 2016); Daragh Murray, *Human Rights Obligations of Non-State Armed Groups* (Hart 2016); Andrew Clapham, *Human Rights Obligations of Non-State Actors* (Oxford University Press 2006); Philip Alston (ed), *Non-State Actors and Human Rights* (Oxford University Press 2005).

25 See e.g. 'Is it Time for a UN Treaty on Violence against Women?' (2018) 22 International Journal of Human Rights 305–24; 'The CEDAW Committee and Gender-Based Violence against Women: General Recommendation No 35' (2017) 6 International Human Rights Law Review 263–78; *The Istanbul Convention, Domestic Violence and Human Rights* (Routledge 2017).

26 See e.g. Dorothée Baumann-Pauly and Justine Nolan, *Business and Human Rights: From Principles to Practice* (Routledge 2016); Surya Deva and David Bilchitz (eds), *Human Rights Obligations of Business: Beyond the Corporate Responsibility to Respect?* (Cambridge University Press 2013).

27 UN's Guiding Principles on Business and Human Rights <www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf>.

28 *R v TRA* [2019] UKSC 51; see the post on this case by Katharine Fortin on the blog of the European Journal of International Law, 20 November 2019 <www.ejiltalk.org/author/katharinefortin/>.

First, many human rights violations are identified and condemned by international bodies, but, because of the lack of an effective enforcement system, the violations go unremedied. This applies even to decisions taken by the European Court of Human Rights. In the latest annual report of the Council of Europe's Committee of Ministers, which oversees the execution of the court's judgments, we read that 'persistent shortcomings in the effective national implementation of the Convention remain a major concern'.[29] Compensation awards remain unpaid, recommended investigations remain unfinished, and changes to laws remain unmade.

There are also problems with the enforceability of conclusions reached by the monitoring committees established under each of the nine UN human rights treaties I have already mentioned. Those committees issue observations and recommendations on self-assessments which states submit to the committees every few years, and they also (for states that have accepted the practice) take decisions on complaints brought to the committees by disaffected residents of states.[30] But apart from *political* pressure, there are no legal mechanisms in place to ensure that the states concerned comply with the committees' recommendations and decisions. There is also a problem in that the remits of the committees often overlap, and the work that they do sometimes repeats that carried out by similar committees established under European human rights treaties. The UN treaties also issue what are called 'General Comments' on the way their treaties should be interpreted: these comments are not always the product of widespread consultation, and at times they lend themselves to the accusation that the committees are over-reaching themselves, thereby damaging their legitimacy.

It is worth noting in passing that the judgments and opinions of the International Criminal Court and the International Court of Justice do occasionally touch upon human rights, but in most cases they do not. Despite the campaigning of eminent human rights activists such as Manfred Nowak in Austria, there is not yet a UN Court of Human Rights, nor is the prospect of one anywhere on the horizon.[31]

The second practical problem relates to the remedies available to victims of human rights abuses committed by non-state actors.

As it is part of my thesis that national criminal and civil law do implicitly but not explicitly protect human rights, it follows that whatever remedy is made available to the victims of breaches of those laws should include some kind of acknowledgment of the human rights element of the crime or civil wrong. Believe it or not, it is already the law in this jurisdiction, as it is in England and Wales, that whenever a court convicts a person of a crime it must give reasons why it has decided not to order the convicted person to pay compensation to the victim of the crime.[32] I could be mistaken, because I could not find any relevant statistics on this, but my strong feeling is that, to borrow from Hamlet, the obligation to consider the ordering of compensation is much more honoured in the breach than in the observance. To the extent that it *is* honoured, I have never heard of a

29 *12th Annual Report of the Committee of Ministers' Supervision of the Execution of Judgments and Decisions of the European Court* (Council of Europe Committee of Ministers 2018) 7.

30 The only exception is the Committee on Migrant Workers: an individual complaints mechanism is pending and will begin operating when 10 states parties have made the relevant declaration under Article 77 of the 1990 Convention on the Rights of Migrant Workers. So far only four states parties have made that declaration.

31 See e.g. M Nowak, 'On the Creation of a World Court of Human Rights' (2012) 7(1) National Taiwan University Law Review 257–91 also at <https://ssrn.com/abstract=2118352>. See too Ignacio de la Rasilla, 'The World Court of Human Rights: Rise, Fall and Revival?' (2019) 19(3) Human Rights Law Review 585–603.

32 Criminal Justice (NI) Order 1994, Art 14(1).

court awarding such compensation in a criminal case because the victim's human rights have been breached, although I cannot see anything in the wording of the legislation to prevent it from being interpreted in that way.[33]

For examples of deficiencies in the practical implementation of our human rights law we need look no further than at the arduous struggles of two groups of victims here in Northern Ireland. The first comprises victims of the widespread child abuse which the recent Hart Inquiry found had occurred in numerous public and private institutions between the foundation of the state and 1995.[34] Fortunately, an Act to provide compensation for those victims obtained Royal Assent just hours before Parliament was dissolved earlier this month[35] and a Redress Board is soon to be established under the chairmanship of Mr Justice Colton. It is good, too, that at least one of the churches which ran the children's institutions has already pledged to contribute to the compensation fund.

The second group of victims consists of those who suffered as a result of supposedly politically motivated crimes committed during the conflict here from 1969 onwards. Proposals for their compensation are out for consultation at the moment.[36] I hope that, amongst others, the many thousands of victims of 'paramilitary-style attacks' – better known as 'punishment shootings and beatings' – will successfully apply for payments under that scheme. It is just regrettable that the assets of the paramilitary organisations in whose names those attacks were carried out can probably not be tapped for contributions to that compensation fund.

My third and final point under the heading of practical problems has to do with the continuing call in some quarters for a Bill of Rights for Northern Ireland. I am not now going into the history of that call, and I acknowledge that there are respected human rights activists who are still very much in favour of the Westminster Parliament enacting such a Bill. My own view is that, while we do need some human rights reforms, we are not currently in a human rights crisis in Northern Ireland. The Human Rights Commission's latest Annual Statement on the human rights situation in Northern Ireland, which is very thorough,[37] does not support the notion that we are in crisis mode, and that was issued *before* the latest reforms relating to same-sex marriage and reproductive rights. I would be surprised if the Commission's Annual Statement next month talks of a crisis.

We do need reforms on several fronts, but focusing on a Bill of Rights is no longer the most efficient or effective way of achieving them. We require much stronger protection of the right to integrated education, removal of the exemption relating to school teachers from the law on fair employment, protection of older people from being discriminated against when accessing services, especially financial services, better guarantees around the use of Irish, and changes to our law on defamation so as to broaden the freedom of speech of journalists and researchers. All of those changes, along with others, can more swiftly be achieved by campaigning for separate pieces of legislation than by waiting for the big bang that an all-embracing Bill of Rights would

33 Compensation can be paid under s 8 of the Human Rights Act 1998, but only in limited circumstances.

34 The Inquiry's report is available at <www.hiainquiry.org/historical-institutional-abuse-inquiry-report-chapters>.

35 Historical Institutional Abuse (NI) Act 2019.

36 'A Legal Framework for a Troubles-Related Incident Victims Payment Scheme' issued by the UK government on 22 October 2019 <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/841467/VP_consultation__FINAL_.pdf>.

37 'The 2018 Annual Statement' <www.nihrc.org/uploads/publications/2018_Annual_Human_Rights_Statement.pdf>.

entail. We have enough political bones of contention in this place without continuing to squabble over one that has really very little meat on it.

Conclusion

In conclusion, let me sum up what I have been trying to argue in this talk, however inadequately. *Firstly*, it is not possible to reach a consensus on the most appropriate way in which to philosophise about human rights, so we ought not to allow disagreement over that issue to stop us from moving forward on the human rights front. *Secondly*, when deciding whether to accept that a novel claim should be given the label of a 'human right', we should look primarily at whether there is yet an international consensus that that claim can be embraced by a right already contained in a UN or Council of Europe treaty. *Thirdly*, if there is local consensus on a claim being recognised as a human right, even though it is not yet recognised at the international level, it should still be protected by law because how else can international consensus change if each jurisdiction does not contribute its own view to the collectivity of views? *Fourthly*, if human rights are values worth protecting, they should be protected not just against violations perpetrated by states but also against violations perpetrated by other organisations and individuals.[38] Equality law is so applied, so why not human rights law? The growing realisation that perpetrators of domestic violence, child abusers, exploitative businesses and non-state armed groups also violate human rights should make the concept more acceptable to those who are unduly sceptical of it in the first place because they perceive it as being confined to abuses committed by the state. *Fifthly*, more should be done at the legal level to ensure that the international system for implementing the recommendations and decisions of monitoring committees and international courts operates more effectively.[39] *Sixthly*, all victims of human rights abuses, especially abuses that leave long-lasting scars, physical or mental, should be entitled to have their status as human rights victims officially recognised and to benefit from more effective remedies, not necessarily confined to monetary compensation. *Finally*, what matters as far as the protection of human rights in domestic law is concerned is not the vehicle which the law uses to confer its protection, but rather the effectiveness of the protection actually granted.

I do not know how many of those conclusions Stephen Livingstone would have agreed with, but I am confident he would have given me a fair hearing. Thank you for doing likewise.

38 In some countries human rights provisions do have what is called 'horizontal' effect. South Africa's 1996 Constitution, for instance, provides in s 8(2) that: 'A provision of the Bill of Rights binds a natural or a juristic person if, and to the extent that, it is applicable, taking into account the nature of the right and the nature of any duty imposed by the right.' And in s 39(2) that: 'When interpreting any legislation, and when developing the common law or customary law, every court, tribunal or forum must promote the spirit, purport and objects of the Bill of Rights.'

39 EU law, for example, allows fines to be imposed on entities, including states, if they breach certain provisions of EU law such as those relating to competition law or the rules on state aid.

NILQ 70(4): 479–501

Barriers to High Court appointments in Northern Ireland

JOHN MORISON,
BRICE DICKSON
AND
ANDREW GODDEN*
Queen's University Belfast

Abstract

This article reports on research carried out on why the Northern Ireland Judicial Appointments Commission is finding it difficult to attract applicants of the appropriate quality for vacant High Court appointments. It makes recommendations for how the appointments system might be changed.

Keywords: High Court appointments; barriers to judicial appointments

1 Introduction

This article examines a significant problem in the legal system of Northern Ireland – the difficulty in recruiting lawyers with the requisite skills to serve as High Court judges. It summarises and slightly expands upon research which two of us were commissioned to conduct for the Northern Ireland Judicial Appointments Commission (NIJAC) in early 2019.[1] The research questions were: what are the real and perceived barriers that may be influencing those at relatively senior levels in the legal professions (widely defined) when they are making decisions about whether to apply for a High Court position and how might those barriers be overcome?

It had become apparent that such positions were not as attractive to potential applicants as in previous years. In its submission to the Review Body on Senior Salaries in 2018, the High Court Judges' Association of Northern Ireland expressed the view that the unsuccessful recruitment round for High Court appointments conducted in 2016 was evidence that the role was not sufficiently rewarding to make legal practitioners want to give up their much higher incomes derived from private practice. The Association

* Respectively, Professor of Jurisprudence, Emeritus Professor of International and Comparative Law and Lecturer in Law, all at the School of Law, Queen's University Belfast. We are very grateful to the two anonymous reviewers who commented on an earlier draft of this article. The views herein expressed, and the responsibility for any errors, remain ours alone.

1 *Barriers to High Court Appointments in Northern Ireland*, submitted to NIJAC in June 2019. The research was assisted by Leah Trainor, a doctoral student in the School of Law at Queen's University Belfast. NIJAC has not yet placed the report on its website <www.nijac.gov.uk>. In its Business Plan for 2019/2020 one of NIJAC's objectives is to 'develop, implement and deliver a programme of action to attract applications [for the judiciary] from the widest possible pool', and its first key milestone in delivering that objective is listed as 'Analysis and dissemination of QUB research findings by March 2020': see <www.nijac.gov.uk/sites/nijac/files/media-files/NIJAC%20Business%20Plan%202019-2020%20Reviewed%20November%202019_0.pdf>.

described the failure to fill three posts out of the statutory complement of 10 as 'little short of disastrous'.[2] Happily, two of the three vacancies were filled in 2018, but one vacancy remained. A competition for that post took place during 2019 but, again, NIJAC felt unable to make an appointment from those who applied for it.

The research into the barriers to High Court appointments in Northern Ireland took full account of existing literature in the field, especially in relation to High Court appointments in England and Wales, where there have been even more serious recruitment difficulties than in Northern Ireland.[3] NIJAC permitted the research to be conducted without any interference or influence, and the views expressed by our interviewees and consultees were cross-checked to ensure that we were being as objective as possible in stating and assessing them. Our findings largely mirror those reached by researchers in England and Wales: dissatisfaction with the pay and pension associated with a High Court position plays a large role in putting off applicants who are already earning more at the Bar or as senior solicitors; the nature of the High Court's role is less appealing than it was because the workload involved is more onerous than ever and the ability to work flexibly is less than in the private sector, or even in public-sector legal bodies such as the Public Prosecution Service; the judicial role continues to be a lonely one, with little assistance provided for research, whereas senior lawyers often greatly enjoy operating in partnership with other lawyers; and the presumption continues to prevail that the applicants best suited to the High Court bench are Queen's Counsel who have spent years impressing judges already on the bench with their advocacy skills, while other attributes required for success as a judge are, relatively speaking, downplayed, a situation which might reflect unconscious bias.

Interviews and consultations were conducted with a large number of interested parties: barristers, solicitors, lawyers working in the public sector, currently serving County Court and High Court judges, and retired County Court and High Court judges. A total of 50 lawyers – 31 male and 19 female – engaged with our fieldwork, 25 through one-to-one interviews and 25 through participation in group discussions. They comprised 18 serving or retired judges (including seven serving High Court judges), 12 solicitors, eight barristers (all QCs), six salaried judges working in tribunals, and six lawyers employed in government legal services. We guaranteed anonymity to all of the interviewees and consultees, but each of them granted us permission to use our discretion in quoting what they said. Gender dimensions, while clearly present and significant, were not a particular focus of this research as they were addressed in earlier work,[4] but many of the structural issues about progression within the legal profession play out at the point where appointment to the High Court is being considered.

In our interviews we tended to focus on the headline issues raised in the Turenne and Bell report from 2018, because many of the challenges facing potential applicants for High Court positions in Northern Ireland are likely to be the same as those arising in England and Wales. Northern Ireland is different, however, in that its smallness as a jurisdiction means that lawyers there are not as able as those in England and Wales to specialise in particular areas of law, to serve as part-time judges while maintaining their

2 Copy of submission on file with the authors.

3 Much of the relevant literature is alluded to in Sophie Turenne and John Bell, *The Attractiveness of Judicial Appointments in the United Kingdom: Report to the Senior Salaries Review Body* (2018), available at <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/748580/SSRB_Report_Attractiveness_Turenne-Bell_Revised_14_March_FINAL_-_temp_pdf.pdf>.

4 See e.g. *Rewarding Merit in Judicial Appointments? A Research Project Undertaken by the School of Law, Queen's University Belfast for the Northern Ireland Judicial Appointments Commission* (2013).

practices without risking a conflict of interest, or to keep confidential the fact that they are applying for a judicial post.

Our five group discussions functioned similarly, with the questions asked of the group being close to those asked of individual interviewees. Only one of the five groups contained serving judges (some salaried, some fee-paid). This was the 'Group of Presiders', that is, judges who have been asked by the Lord Chief Justice to preside over the group of courts or tribunals within which they are a judge, such as the County Courts, the magistrates' courts, the industrial tribunals (called employment tribunals in England and Wales), appeal tribunals (for welfare benefit claims) and the criminal injuries compensation appeals panel. The other four groups comprised solicitors (two groups), lawyers working for the Public Prosecution Service, and staff from the Crown Solicitor's Office. The opinions expressed during the group discussions were added to those expressed by individual interviewees and all were then assessed as a whole.

The High Court of Northern Ireland

As in England and Wales, the High Court of Northern Ireland constitutes the highest level court for first instance hearings in the jurisdiction. Its powers are set out in sections 16 to 33A of the Judicature (NI) Act 1978, as amended. Since 2004 the currently permitted number of High Court judges is 10, not including the Lord Chief Justice or the three Lords Justices of Appeal.[5] At present only eight of the 10 High Court positions are filled.[6] There are currently two women judges in the High Court and a further recently appointed judge was previously a solicitor. The vast bulk of High Court work is in the area of civil law, although High Court judges do occasionally sit in the Crown Court to hear serious criminal cases, and in the High Court they hear judicial review applications relating to criminal causes or matters. Unlike their High Court colleagues in England and Wales, however, they do not hear appeals from criminal cases in magistrates' courts, a task which is assigned instead to County Court judges, who also handle many other serious criminal cases.[7] Occasionally, High Court judges will also sit in the Court of Appeal of Northern Ireland. Between them they are called upon to perform more than 60 other tasks at the Lord Chief Justice's request, ranging from being a Judge-in-Residence at Ulster University or Queen's University Belfast (positions which carry relatively light duties) to chairing bodies such as the Council of Legal Education or the Boundary Commission for Northern Ireland (which at certain times can be onerous responsibilities). There is a particular problem with so-called 'legacy' inquests relating to deaths which occurred during the conflict in Northern Ireland between 1969 and 1998. To help solve this problem the Lord Chief Justice has been appointed President of the Coroners' Court, Mrs Justice Keegan has been appointed head of the Coroners Service[8]

5 S 2(1) of the Judicature (NI) Act 1978, as amended by Art 2 of the Maximum Number of Judges (NI) Order 2004 (SI 1985), provides: 'The High Court shall consist of the Lord Chief Justice of Northern Ireland (in this Act referred to as "the Lord Chief Justice") and not more than ten puisne judges who shall be styled "Judges of the High Court".'

6 In addition to the vacancy which was competed for but not filled in 2019, a further vacancy was created when Sir Bernard McCloskey was promoted from the High Court to the Court of Appeal in September 2019.

7 For more on the role of County Court judges see pp 494–496 below.

8 For more details, see <www.justice-ni.gov.uk/articles/coroners-service-northern-ireland>.

and individual High Court judges have at times been asked to conduct particularly controversial inquests.[9]

Applicants for vacancies on the High Court are recommended for appointment by NIJAC after responding to an advertisement and undergoing a rigorous selection process. NIJAC was established in 2005 under the Justice (NI) Act 2002, amended by the Justice (NI) Act 2004. In the years between 2008 and 2018 it organised seven competitions for High Court appointments.[10] In total there were 55 applicants for the posts, of whom 20 were shortlisted, and 10 were eventually offered a post. Of these 10, nine were Queen's Counsel and one was a senior solicitor. No County Court judge has been appointed to the High Court since NIJAC was established, even though there have been a number of applications from County Court judges. In this context it is worth bearing in mind the words of the Lord Chief Justice of England and Wales, Lord Burnett, in his Treasurer's Lecture delivered on 18 February 2019: 'if the judiciary is not appointed from every corner of the legal professions, talented people will be missed and the overall quality of the judiciary will suffer'.[11]

The findings of research already conducted

The supposed 'recruitment crisis' presently facing the judiciary in general has attracted some recent attention.[12] We already know quite a lot about its nature and possible causes. In the UK work on the theme began in 1998,[13] and the most recent study was completed in 2018 by Turenne and Bell.[14] This builds upon findings in annual Judicial Attitudes Surveys from 2014 and 2016,[15] continuing interest from the Review Body on Senior Salaries[16] and work done by the House of Lords Select Committee on the Constitution.[17]

9 E.g. into the 'Ballymurphy massacre', when 10 civilians were allegedly killed by British soldiers in the 72-hour period following the introduction of internment on 9 August 1971: see <https://judiciaryni.uk/ballymurphy-inquest>. Keegan J is due to give her verdict in that inquest during the first half of 2020. In November 2019 she announced a provisional timetable for the completion of the remaining 38 legacy inquests, some of which involve multiple deaths: see <https://judiciaryni.uk/sites/judiciary/files/media-files/Presiding%20Coroner%27s%20Statement%20in%20relation%20to%20legacy%20inquests%20-%2020%20Nov%202019.pdf>.

10 By the end of 2019 NIJAC had not disclosed how many applicants there were for the 2019 competition.

11 'A Changing Judiciary in a Modern Age', 5, available at <www.judiciary.uk/wp-content/uploads/2019/02/mt-treasurers-lecture-final-for-publishing.pdf>.

12 See, for example, Adrian Jack, 'A Low Benchmark?' *New Law Journal* (London, 12 January 2017) 6; 'The Recruitment Crisis is Damaging Every Level of the Judiciary' *The Times* (London, 8 March 2018); Michael Beloff, 'Judge Not: Few QCs these Days Aspire to the Bench', *The Spectator* (London, 25 June 2019). See also Judiciary of England and Wales, *The Lord Chief Justice's Report* (2017).

13 Hazel Genn, *The Attractiveness of Senior Judicial Appointment to Highly Qualified Practitioners* (Judicial Executive Board 2008) <www.ucl.ac.uk/judicial-institute/sites/judicial-institute/files/the_attractiveness_of_senior_judicial_appointment_research_report.pdf>.

14 See n 4 above.

15 See the two reports by UCL's Judicial Institute at <www.judiciary.uk/wp-content/uploads/2015/02/jac-2014-results.pdf> and <www.judiciary.uk/wp-content/uploads/2017/02/jas-2016-england-wales-court-uk-tribunals-7-february-2017.pdf>.

16 See the Senior Salaries Review Board, *Thirty Ninth Annual Report on Senior Salaries 2017, Report 87* (CM 9455, 2017) and its follow-up work in the NatCen, *Survey of Newly Appointed Judges in the UK in 2017* (2018) <www.gov.uk/government/publications/natcen-report-on-survey-of-newly-appointed-judges-in-the-uk-2017>.

17 House of Lords Select Committee on the Constitution, *Judicial Appointments: Follow-up* (HL Paper 32, 2017–2019) <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/629679/SSRB_2017_report_Print.pdf>.

In Northern Ireland, the general position on judicial appointments has been fairly well researched.[18] Although the research discussed in this paper is the first study looking at the specific issue of High Court appointments in Northern Ireland, there is some information that can be read across from the work carried out for the UK as a whole, as well as a number of internal reports that can be read alongside NIJAC's reports.[19] The existing reports revealed a number of themes which were explored further during the research which is now being presented.

2 Pay and pensions as barriers

Amongst the commonest barriers to senior judicial appointments mentioned to us by our interviewees were: (a) the relatively low salary attached to the position of High Court judge in comparison with the higher salaries earned by successful QCs and senior partners in firms of solicitors; and (b) the changes applied to the judicial pension scheme in 2015.

Pay

In the financial year 2018–2019 the salary of a puisne High Court judge in Northern Ireland (as in England and Wales) was £185,197. From 1 April 2019 it has been £188,901.[20] The Senior Salaries Review Board confirmed in 2018 that judicial salaries had been allowed to fall far below the levels they should have been at, and it therefore recommended that the salary of a High Court judge should be raised by 32 per cent to £240,000, backdated to April 2018. On 26 October 2018 the Lord Chancellor announced that judicial pay for 2018–2019 would rise by just 2 per cent – still the biggest rise in 10 years – and in June 2019 the Ministry of Justice introduced a new temporary *allowance* for High Court judges in England and Wales in order to improve retention and recruitment.[21] The new allowance is to be backdated to 1 April 2019 and has been set at 25 per cent of a High Court judge's basic salary, replacing the previous allowance of 11 per cent which was introduced in 2017. It applies only to High Court judges who are already, or will become, eligible for the new judicial pension scheme introduced in 2015, and it will remain in place only until a sustainable long-term solution is introduced to deal with the pension problems (on which see the next section of this article).[22] The Ministry of Justice is working with the devolved administrations in Northern Ireland and Scotland as regards

18 Queen's University Belfast, *Propensity to Apply for Judicial Office under the New Northern Ireland Judicial Appointments System: A Qualitative Study for the Northern Ireland Judicial Appointments Commission* (2008) and *Rewarding Merit in Judicial Appointments? A Research Project Undertaken by the School of Law, Queen's University Belfast for the Northern Ireland Judicial Appointments Commission* (2013) and further commentary and analysis of these reports in J Morison, 'Finding Merit in Judicial Appointments' in C Dwyer and A-M McAlinden (eds), *Justice in Transition* (Hart 2015) and in J Morison, 'Beyond Merit: The New Challenge for Judicial Appointments' in G Gee and E Rackley (eds), *Debating Judicial Appointments in an Age of Diversity* (Routledge 2017) 223–239.

19 B Coulter and E Doyle, *Review of the Effectiveness of NIJAC's Appointments Process* (2018); Lord Chief Justice's Office, *Results of Survey of QCs on Barriers to Seeking Judicial Appointment* (2018) and NIJAC's *Annual Reports and Accounts* dating from 2006 <www.nijac.gov.uk/publications/type/annualReport?search_api_views_fulltext=annual%20&action=search>.

20 See the announcement by the Ministry of Justice at 3 <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/836749/judicial-salary-schedule-oct-2019.pdf>.

21 For details, see *Government Response to Report No 90 of the SSRB: Major Review of the Judicial Salary Structure* (5 June 2019, CP 107) <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/806480/government-response-ssrb-june-2019.PDF>. See too Owen Boycott, 'High Court Judges Get Pay Rise to Deal with Hiring Crisis in England and Wales' *The Guardian* (London, 5 June 2019) <www.theguardian.com/uk-news/2019/jun/05/uk-high-court-judges-to-be-given-25-pay-rise-to-tackle-hiring-crisis>.

22 Ibid paras 19–22.

the changes that may be made in those jurisdictions, and further details are to be announced in due course.[23]

Although the average earnings of QCs in Northern Ireland during that or any other year is not known, it is an accepted fact that many QCs regularly earn much more than the salary of a High Court judge each year. As long ago as 2012 it was disclosed in an answer to a question put to the Minister for Justice in the Northern Ireland Executive by Jim Allister MLA that in the five previous financial years 21 Queen's Counsel had each earned more than £1,000,000 in legal aid payments alone and £55,000,000 was awarded to 70 QCs in all.[24] Although legal aid payments have been reduced since 2012, considerable rewards are still available to barristers and solicitors through non-publicly funded work. As far as solicitors are concerned, the head of the Belfast branch of an international firm asked us rhetorically why he or she would want to take a 50 per cent pay cut in order to become a High Court judge.

It seems clear, certainly, that in Northern Ireland there is a cadre of high-earning lawyers who, at present, are not likely to be interested in applying to become a High Court judge because they would be significantly better off financially if they stayed in their current job. As discussed in section 3 of this paper, they may be individuals who find their current work much more enjoyable than the work they anticipate having to do as a High Court Judge. We know from Turenne and Bell's report that the same situation obtains in England and Wales,[25] although those authors also noted that for solicitors in Northern Ireland the judicial salary is 'often close enough to what would be received in private practice'.[26] They said the same about solicitors in Scotland but also cited a survey by the Law Society of Scotland which indicated that, in 2017, at least 25 per cent of all solicitors, regardless of the size of the firm where they were working, earned more than £200,000 per year.[27]

It is equally clear that pay, in and of itself, is not a barrier to High Court appointment as far as serving County Court judges in Northern Ireland are concerned. In 2018–2019 the salary of a County Court judge was £148,527[28] (with the Recorder of Belfast receiving an 8 per cent uplift), meaning that, for applicants from that pool, promotion to the High Court would entail a pay rise of almost 25 per cent. Lawyers working in the public sector – for example, in the Public Prosecution Service or the Crown Solicitors Office – would also be earning much less than the salary of a High Court judge, indeed less than that of a County Court judge too.

23 Ibid para 23.

24 Adrian Rutherford, 'Revealed: The QCs and their Legal Aid Millions' *Belfast Telegraph* (Belfast, 5 June 2012) available at <www.belfasttelegraph.co.uk/news/northern-ireland/revealed-the-qcs-and-their-legal-aid-millions-28756954.html>. See too the *Belfast Telegraph* of 25 June 2013 <www.belfasttelegraph.co.uk/news/northern-ireland/lawyers-paid-100m-in-legal-aid-29371471.html>, 23 December 2013 <www.belfasttelegraph.co.uk/news/northern-ireland/top-legal-aid-barrister-earned-1m-29862020.html> and 24 December 2013 <www.belfasttelegraph.co.uk/news/northern-ireland/northern-ireland-90m-legal-aid-gravy-train-not-sustainable-says-mla-29863234.html>.

25 Turenne and Bell (n 3) paras 32–38.

26 Ibid para 33.

27 Financial Benchmarking Survey Overview <www.services.tribalgroup.com/apps/LSS_Reports>.

28 This pay is slightly higher than that earned by Circuit judges in England and Wales because of the fact that County Court judges in Northern Ireland may be required to preside in serious criminal cases without the assistance of a jury (in terrorism-related cases). For 2019–2020 the salary of a County Court judge in Northern Ireland was raised to £151,498: see n 20 at 4.

Pensions

The pension arrangements for all judges in Northern Ireland, as in the rest of the UK, were radically altered by the Judicial Pension Regulations 2015. The changes were described to us by one judge as the 'biggest obstacle for the traditional applicants . . . for both silks and partners in large successful commercial practices'. The 2015 Regulations affected judges in three different ways depending on their length of service and their date of birth:

1. those who were members of the judicial pension scheme before 1 April 2012 and were born before 2 April 1957 retained full protection;
2. those who were members of the scheme before 1 April 2012 and were born between 2 April 1957 and 1 September 1960 were entitled to tapering protection until a date between 31 May 2015 and 31 January 2022, whereupon they are to be excluded from the scheme and become entitled to join the new scheme (they also acquired the option to transfer to the new scheme on 1 April 2015);
3. those who were members of the scheme before 1 April 2012 but were born after 1 September 1960 were not entitled to any protection and were excluded from membership of the scheme after 1 April 2015, on which date they were able to join the new scheme.

It follows that judges who fell within the third category were treated less favourably than those who fell within the first and second categories, and those who fell within the second category were treated less favourably than those who fell within the first category. The factor determining which category a judge fell into was their date of birth. In December 2018 the Court of Appeal of England and Wales confirmed the decision of the Employment Appeal Tribunal (and of the employment tribunal below that) that the changes to the pension scheme discriminated against certain judges on the ground of their age and that the Lord Chancellor and Ministry of Justice had failed to show that such discriminatory treatment was a proportionate means of achieving a legitimate aim.[29] In June 2019 the UK Supreme Court refused leave to appeal this decision, which means that the UK government may be required to compensate those judges who have been unfairly discriminated against on age grounds. As things stand, however, judges who were not members of the judicial pension scheme before 1 April 2012 will apparently not qualify for any such compensation. We were told that complaints about age discrimination in the implementation of the new judicial pension scheme had also been made by various judges in Northern Ireland and that the determination of those complaints had been put on hold until the final outcome of the litigation in England and Wales was known. Whatever remedy is eventually put in place for the victims of the age discrimination there, it is likely that it will be replicated in Northern Ireland too in due course.

It is clear that some High Court judges, particularly the younger ones, now have pension arrangements which are not as generous as those previously in place. But in that respect those judges are no different from many other categories of workers in the public sector whose final-salary pension schemes have recently been abolished or radically altered and whose predecessors in their jobs were treated more favourably. Moreover, we heard from several interviewees that successful barristers, like other self-employed individuals and practising solicitors, usually build up their own private pension pot over a number of years

29 The *Lord Chancellor v McCloud* [2018] EWCA Civ 2844 (20 December 2018). The details of the schemes are set out in paras 12–24 of the Court of Appeal's judgment.

through savings and investments. If they become senior judges, they will still enjoy the benefits of that pension pot in addition to the pension which they will receive under the new judicial pension scheme when they retire from the bench. In addition, the pension problem is a transient one because, as time goes, by applicants for senior judicial appointments will no longer feel aggrieved that the previous very generous pension scheme is no longer available to new appointees. Changes to tax rates, moreover, are a vicissitude which all workers throughout the country have to face from time to time, and there appears to be no justification for treating members of the judiciary as a special case.[30]

In short, while recognising that the change in pension arrangements has affected some serving judges quite adversely, it is a phenomenon which affects all tiers of the judiciary, as well as many other public-sector workforces, and it is unlikely to be a discernible barrier to lawyers who are considering whether to apply for future senior judicial posts.

3 The nature of the job as a barrier

Research conducted for the NatCen Survey of Newly Appointed Judges suggests that those who join the judiciary are motivated by the idea of undertaking a challenging job, providing a valuable public service and taking what may be seen as a natural career step. Moreover, these expectations are generally fulfilled.[31] Turenne and Bell list the main incentives for applying for appointment as a reduction of workload and pressure compared to private practice, a secure salary (albeit often smaller), a good pension, respected social status and a wish to put something back into the legal system. Our interviewees in Northern Ireland confirmed these incentives to a degree, but they also pointed to a number of changes to the way (barristers and solicitors) work and to the workload of judges, the combined effect of which has had an adverse impact on the attractiveness of a judicial role.

Changes to the job of legal practitioners

The life of the practitioner has changed in recent decades. We were told, for example, that 20 or so years ago there was both more work available and considerably more pressure to take what was on offer. In this context, despite the reduced salary, a judicial appointment had a particular appeal. As one current judge told us, 'in the past the bench was seen as an escape from the unrelenting pressure of work, one case after another'. Nowadays, the nature of practice is said to have changed. The same judge pointed out that:

> . . . now the vast majority [of barristers] have a much better work/life balance . . . [for me] there was no question of taking time off to go skiing . . . [but] now you can organise your cases . . . there is less work available . . . [but] you can still have a decent life.

Several of our female interviewees particularly stressed the importance of this flexibility.

Both barristers and solicitors told us that now even very successful practitioners have considerably more control over their work. Indeed, a senior partner in a solicitors' firm referred to the hierarchy and formality of the 'back corridor' (the area in the Royal Courts of Justice building where High Court judges have their chambers) as comparing unfavourably to the collaborative and teamwork approach adopted in the modern

30 Judges in Ireland strongly campaigned against cuts to their pay and pensions after emergency measures were taken to reduce public expenditure in the wake of the financial crisis in that country in 2008. Despite dissatisfaction on the part of many serving judges, the changes do not appear to have had much of an effect on the willingness of lawyers to accept senior judicial appointments since that date.

31 Above n 8.

solicitors' firm, where diversity and flexibility in working patterns are valued for business reasons. Similarly, as a senior barrister put it:

> I am going to have to come in every day of the week. I am going to have a boss . . . who will tell me where I have to go and what cases I have to hear . . . and my holidays are going to be cut back . . . and all this for a cut in pay.

To be fair, we heard contrasting views as well. One judge with experience of the back corridor thought it was 'very collegiate', and several such judges told us they could easily informally consult one of their colleagues on a legal point if they felt the need to do so. Another judge appreciated the variety and challenges which High Court work presented.

It seems, nevertheless, that some modernisation of the working conditions of judges might be worth consideration. In particular, as one QC put it, 'if there was the option of having part-time hours that would make it much more attractive'. While this is problematic in terms of the statutory provisions setting out the establishment of the judiciary, it is an important suggestion to consider for the future. Two judges told us, however, that it would not make sense to appoint a part-time judge unless he or she could commit to being available for a two- or three-month period at a time and thereby able to take on long cases. Similarly, there was support from at least three serving High Court judges for the idea of targeting recruitment at senior practitioners, perhaps in their sixties, on either a full-time or part-time basis, and perhaps with an extension to their retirement age.[32] We return to the issue of part-time and flexible working below.[33]

Changes to the job of High Court judges

At the same time as the work of practitioners has become more flexible, the work level required of High Court judges has intensified. Several interviewees mentioned that the volume of work is much greater now than it was 20 years ago. Most judges accept this. As one said:

> . . . there are a lot of good things about being on the bench and most people broadly enjoy it . . . [W]hen I applied I did expect that I would have to work pretty hard . . . At times you would feel a little stressed but no more than at the Bar . . . we shouldn't be too defensive.

Indeed, all of the serving and retired High Court judges we interviewed told us that the experience of serving as a High Court judge was enjoyable, and none expressed any regrets at having taken up the role.

However, legacy cases relating to the troubles in Northern Ireland (e.g. inquests, compensation claims and judicial reviews of actions or inactions by the police) and historic sex abuse cases were seen as particularly difficult and unattractive. One judge spoke eloquently about the difficulties he has had as an older person in going back into the past to revisit former times in order to try to apply modern standards to old cases. We were also told that the current political situation in Northern Ireland, where the Executive and Assembly have not met since March 2017, has put additional pressure on judges who are finding that issues are being passed to the judiciary in the absence of any other decision-making forum. This arises from the fact that the Northern Ireland (Executive Formation and Exercise of Functions) Act 2018 leaves it uncertain as to what decisions

32 See further comments made by Lady Hale to the House of Lords Constitution Committee, reported in *The Times*, 28 March 2019 <www.thetimes.co.uk/article/let-judges-stay-on-until-72-urges-hale-k8txk3mts>.

33 See pp 497–498 below.

can or cannot be taken by civil servants in the absence of a serving minister.[34] While concerns about security are no longer very prevalent, the compensations of the status of the post, including the knighthood or damehood, are perhaps no longer sufficient to overcome the issues about isolation and the effect on family life that still persist.

Concerns were expressed too about the variety of specialisms that might be involved in the High Court, some of which might be far from an applicant's expertise in the increasingly specialist world of legal practice. We were told that, 'by the time you got to the stage in your career where you were ready to apply to a High Court job . . . you will be a specialist and then you then have to go back to being a generalist'. As one QC put it, 'You may find that you're pitched in somewhere . . . like the Family Division where you never had any experience or interest and you can do very little about it.' This might be an argument for increased training of new judges or, alternatively, for ensuring that judges are appointed more directly to particular specialisms rather than rotating across areas of law. Such changes might assist in widening those attracted to making an application. One interviewee told us, 'I believe that it would encourage solicitors in particular who have a speciality . . . I am a criminal lawyer, what do I know about Chancery cases?' At the same time a serving judge commented that one of the attractions of the post was 'the new challenge . . . You are being asked to do things that are difficult . . . and moving into a different area of law.'

The effects of austerity too are being felt, reinforcing views found in the wider literature about judges' working conditions.[35] While we did not find the same complaints about facilities in the High Court of Northern Ireland as others have reported in Great Britain, there is a sense in which the funding environment has sapped morale. As one judge put it: 'austerity has been a factor . . . ideas which could help judges be more efficient are not viewed favourably at this time . . . there is a lot of scope here . . . [for] someone to do research and work with and provide a critical perspective'. Indeed, the idea of judicial assistants was popular both with serving judges and with some potential applicants. As one judge observed, confirming the considerable support amongst nearly all of the serving High Court judges we interviewed for an enhanced system of judicial assistants, 'it would really help to have someone . . . setting out the background, explaining, for example, the contractual framework in a complex case . . . or doing a bit of research on the caselaw'. Another linked this to the increase in personal litigants, remarking that: 'I get nothing from them [in terms of legal argument] . . . and end up having to go and do my own research.'

This is an area in which reforms could be introduced to considerable effect. A small research unit already exists in the Royal Courts of Justice, but in practice it is available for use only by judges who are sitting in the Court of Appeal. While there is a system of tipstaffs who attend judges, they are not really involved in legal work. Securing more

34 The Act was passed as a consequence of the decision of the Court of Appeal of Northern Ireland in *In re Buick* [2018] NICA 26, where the court held that, in the absence of a Minister, the relevant government department did not have the power to make a decision to grant planning permission for a major waste incinerator. Section 3(1) of the Act provides: 'The absence of Northern Ireland Ministers does not prevent a senior officer of a Northern Ireland department from exercising a function of the department during the period for forming an Executive if the officer is satisfied that it is in the public interest to exercise the function during that period.' When the Attorney-General of Northern Ireland referred the issue of the extent of departmental powers to the Supreme Court, the Justices adjourned the reference because it preferred to determine such an issue against the background of a clear factual matrix: *Reference by the Attorney General for Northern Ireland of devolution issues to the Supreme Court pursuant to Paragraph 34 of Schedule 10 to the Northern Ireland Act 1998 (No 2)* [2019] UKSC 1.

35 See Turenne and Bell (n 3) 19.

funding for proper support staff for judges would be a development that would make the judicial role more attractive to potential applicants. High Court judges appointed in the Republic of Ireland are now assigned a judicial assistant rather than a tipstaff, although there is apparently some dissatisfaction with the system.[36] Judicial assistants also play a significant role in the judicial apparatus of other jurisdictions; in particular, they are given very high esteem in the Netherlands.[37]

4 The recruitment process as a barrier

General impressions

Overall the recruitment process is not a major problem. As one QC told us, 'I don't see any barriers as such . . . I just wouldn't be interested in the job.' However, several interviewees did tell us that some things were amiss with the recruitment process. One said that 'the application process is not terribly effective', adding that 'proof of [judicial] experience should count as *proof* of ability'. Another said that 'the system itself is irredeemably flawed . . . because there is a [recruitment] panel of four, two of whom are lay people'; the same interviewee described this process as 'not efficient'. A judicial interviewee observed that 'the current selection system is incredibly mathematical and moderated marks cause problems too . . . Obviously the process of appointment is putting off people.' Despite all these comments, most interviewees were otherwise positive about the role of NIJAC in running the competitions for High Court positions.

Demonstrating abilities, qualities and skills

It was frequently suggested to us that barristers were disincentivised by the fact that they had to complete an application form in which they were expected to provide evidence for how they had acquired certain abilities, qualities and skills. Being self-employed, many barristers will rarely if ever have had to complete such a form in the past. This was considered to be in contrast to lawyers working in the public sector, who routinely apply for higher-grade posts and in the process are tested on a range of competencies.

We can see nothing out of place in the list of abilities, qualities and skills which applicants for High Court positions are expected to possess, and it is appropriate that in relation to each of the five abilities, qualities and skills that are listed on the application form applicants are required to provide evidence demonstrating, through the use of examples from their own experience, how they meet the characteristics set out in the three or four bullet points under each of the five headings.

The Applicant Information Booklet informs applicants that the distribution of marks in the assessment is as follows: 40 per cent for intellectual capacity, knowledge and expertise; 15 per cent for personal qualities; 15 per cent for understanding and fairness; 15 per cent for communication skills; and 15 per cent for leadership and management skills. This seems to us to be a reasonable allocation of marks, but we suggest that it might be worthwhile further disaggregating the 40 per cent allocated to intellectual capacity, knowledge and expertise into smaller amounts, such as 10 per cent for intellectual capacity (always very difficult to assess), 15 per cent for knowledge and 15 per cent for relevant expertise. Amongst other benefits, this might allow previous judicial

36 Conor Gallagher, 'Judges Unhappy with Assistants' Support and Want Ushers Back' *Irish Times* (Dublin, 24 November 2017) <www.irishtimes.com/news/crime-and-law/judges-unhappy-with-assistants-support-and-want-ushers-back-1.3303136>.

37 Nina Holvast, 'Considering the Consequences of Increased Reliance on Judicial Assistants. A Case Study on Dutch Courts' (2014) 21 International Journal of the Legal Profession 1.

experience to be given some explicit weight in the selection process. We have been told that this is now under review.

The submission of written work

Applicants are also required to submit three pieces of 'written work of substance'. NIJAC's website contains a cover page for whatever is submitted. This page makes it clear that the three pieces should demonstrate the applicant's 'abilities, qualities and skills . . . and how these could potentially transfer to [the work] expected of a High Court judge'.

The written work that must be submitted by applicants as evidence of their writing ability was seen as problematic by some of our interviewees, both those who had succeeded in being appointed to a judicial role and those who had not. There was a feeling that NIJAC's Selection Committees were more impressed by examples of senior counsel's opinions than by, say, a judgment issued by a County Court judge or academic material (even in book form) submitted by a lawyer in public service. We were told by one interviewee that a County Court judge who applied for a High Court post did not have any written judgment available to submit and that another such applicant submitted a judgment that was flawed. One County Court judge was emphatic that the nature of their work was not suitable for producing polished judgments which might impress a selection panel. The same judge told us that 'the barrister's opinion is very much part of their day-to-day job in the way that it is not for us' and that, while County Court judges could of course write up a 'vanity judgment' with a view to advancing themselves, the pressure of day-to-day working life was such that they would need to make a special effort to focus on doing so as part of an application process.

Several of our solicitor interviewees felt that may have the evidence of high-quality written work but complained that they may have 'a chip on [their] shoulders . . . [W]e don't recognise the evidence that exists . . . [H]ow can you compare a commercial lease . . . with maybe a very good use of precedent . . . a "must" instead of a "shall". . . against an opinion setting out the salient points . . . ?'

Interviewees employed as government lawyers told us that it was difficult for them to submit written material which met the stated criteria because the documents they have written are not theirs to share – they are owned by the parties to the cases in question. In any event, many documents they generate are 'joint efforts', and it can be almost impossible for an applicant from that sector to identify specific parts of the documents which were written exclusively by them.

Generally speaking, concerns were expressed by several interviewees over the fairness of this element of the selection process, especially as the Applicant Information Booklet and the cover page, despite being quite wordy on the matter, do not specify precisely what qualities are being looked for in the written submissions. Nor is there any word limit to what can be submitted and, on occasions, the assessment of submissions must be akin to comparing apples and oranges. As one interviewee remarked: 'What is the purpose behind asking for this? Is this something that could be properly, fairly and more consistently approached as part of the actual assessment process? . . . You could [instead] be asked to read something and write up a brief summary judgment.'

We were left with the distinct impression that more could be done to reduce applicants' understandable anxieties relating to this element of the selection process and that more could be tested at this stage. We also discussed with interviewees whether in this context 'merit' might mean more than straightforward intellectual ability and encompass an ability to run a court decisively and efficiently. This might provide an

opportunity for all applicants to come away from the recruitment process with an equal sense of its fairness. While it may be difficult to produce metrics to demonstrate such additional attributes, the general opinion amongst our interviewees (especially serving and retired County Court judges) was that there are important abilities, qualities and skills that are not presently assessed and, in particular, previous judicial experience is not given sufficient consideration.

Consultees and referees

It was noted by several of our interlocutors that NIJAC's abandonment of 'consultees' during recruitment competitions was a good thing. It means that applicants who are close to existing judges cannot capitalise upon that coincidence in order to boost their application. But one High Court judge was of the very definite view that it was a mistake not to make use of consultees, saying 'I know who is High Court material and who is not.' Another High Court judge suggested that all existing High Court judges should be asked for their views on the candidates being considered by NIJAC (as was supposedly the practice in the past), and another said that the removal of this practice represented a diminution of their status. Our own view is that such consultations should not take place: provided that NIJAC's selection process is rigorous in obtaining as much information as possible about the knowledge and competencies of the applicants (together with honest references), a level-playing field is maintained and the risk of discriminatory treatment or unconscious bias around the nature of what constitutes 'merit' is reduced.

Reliance on references, particularly if the referee is asked to comment specifically on key abilities, qualities and skills in which good applicants are expected to excel, helps to preserve an equal opportunities approach to recruitment. But it was pointed out to us by one senior counsel that non-advocate applicants for the High Court may have difficulty in finding a serving judge who knows their work well enough to provide an impressive reference. Of course, there is no requirement to have a serving judge act as a referee, but doubtless a positive reference from such a person would carry considerable weight. The same source said that the fact that so few solicitors are appointed to the bench was ironic because, on the whole, solicitors are better than barristers at managing cases. But we acknowledge that senior solicitors who apply for a position in the High Court are very likely to know other senior people not just in the law but in other walks of life and therefore should have little difficulty in calling upon them to act as referees.

The key point here is that Selection Committees should not attribute undue weight to a reference just because it emanates from a senior judge. We understand that references are not given 'marks' as such, but they are taken into account by the Selection Committees when they are considering all the other evidence they have collected about the candidates' abilities, qualities and skills during the appointments process. The purpose of the reference is to allow the Committees to confirm the accuracy of information which the Selection Committees already possess.

Personal knowledge of applicants

The interviewees we spoke to differed greatly as regards the appropriateness of members of Selection Committees, taking into account their own personal knowledge of applicants when assessing them. One very experienced judge said it was 'ludicrous' not to allow such personal knowledge to be used, although that judge did not go so far as to suggest that the panel member could fill in the blanks of an acquaintance's poorly completed application form. The same judge thought that Selection Committees should consist only of people who have themselves had judicial experience. But nearly every other

interviewee who was asked about the selection process was of the view that lay people should be involved at all stages, and that the chairperson did not have to be a judge (in Northern Ireland the chairperson has always been a judge).

Barrister interviewees tended to accept that an applicant could be at a distinct advantage if the judges on the Selection Committee happened to be very familiar with his or her work (while acknowledging that this could be a double-edged sword).[38] One senior counsel felt strongly that the selectors of judges did need to know the people they were appointing but, when pressed, the lawyer in question could not provide a justification for that view. Another QC said the main point that he or she would like to get across in their interview was that the selection system for High Court judges should be more subjective than at present: the current system was described as too much like a box-ticking exercise.

We were surprised that during our interviews with some judges the names of particular barristers were occasionally mentioned as people who either should have already applied for High Court posts or should be thinking of doing so in the near future. It seems that certain barristers are 'identified' as potential future High Court judges in a way which does not apply to applicants from other pools, such as solicitors, lawyers in public service or even existing County Court judges, all of whom are not so often within the gaze of High Court judges in the courts. Specific barristers are often strongly encouraged to apply by senior serving judges, which is a dangerous course of action because it may lead to unreasonable expectations on the part of such barristers who do apply. Just as canvassing for particular jobs is sometimes stated to be a justification for disqualifying the canvasser from consideration for the job, so those who are selecting people for particular jobs should not be suggesting in any way that a person would be a worthy candidate for the post. Pre-judging a potential applicant in this way is potentially discriminatory against other applicants.

In a small jurisdiction such as Northern Ireland it is unrealistic to expect people on Selection Committees not to be personally acquainted with some of the applicants for judicial posts. However, good practice would dictate not only that such acquaintanceships are notified to the whole Committee in advance of any deliberations taking place but also that the member who knows an applicant does not make any comments about that applicant to other members of the Committee which are not based on something stated by the applicant in his or her application form or during the course of his or her participation in any role-playing exercise or interview.

We were told that NIJAC's conflict of interest policy is based on guidance issued by the Northern Ireland Audit Office,[39] and this seems appropriate. What is equally important, however, is that the guidance is strictly adhered to. It should be remembered that, according to the guidance, '[a] perception of a conflict of interest can be just as significant as an actual conflict of interest. The key issue is whether there is a risk that a fair-minded outside observer, acting reasonably, would conclude that there is a real possibility of bias.'[40]

There is one additional element relating to the appointment process that is particularly eye-catching. Reflecting on the changes to the appointment process brought about by the advent of NIJAC, and its part in demystifying the process and the role of judge, one

38 This notion of 'visibility' was also flagged up in the 2013 study from Queen's University Belfast on the concept of 'merit' (n 18).

39 Presumably, this is *Conflicts of Interest: A Good Practice Guide* (2015) <www.niauditoffice.gov.uk/sites/niao/files/media-files/conflicts_of_interest_good_practice_guide.pdf>.

40 Ibid para 2.4.

interviewee told us how the 'tap on the shoulder' recruitment scheme perhaps carried with it 'a sense that it was your duty to do it . . . to put something back into the profession . . . and that seems to have gone now . . . it is a job now'.

5 Traditions and assumptions as barriers

The assumed superior suitability of senior counsel

As stated in the introduction to this article, the most usual applicants for High Court appointments during the last decade have been barristers, by far the commonest applicants shortlisted for appointment have been barristers, and the overwhelming majority of those actually appointed have been Queen's Counsel of several years' experience at that level.

Several of our interviewees, including current High Court judges, seemed to take it for granted that senior counsel would be the most likely category of person to merit appointment to the High Court. In particular, Senior Crown Counsel were thought to be particularly eligible with one informant saying 'anyone who was Senior Crown Counsel normally felt that they had to go on to the bench . . . some have enjoyed it . . . one or two of them maybe should not [have made the move]'. When pressed, however, interviewees were not able to provide hard evidence as to why Crown Counsel status or success at the Bar, which tends to be measured in terms of years of experience and perceived income levels, should necessarily mean that those individuals would be better equipped than other applicants to serve as High Court judges. More commonly, we received the impression from judicial interviewees that serving judges who themselves had been successful as senior barristers tended to assume that the competencies displayed in that latter role by the subsequent generation of senior barristers inevitably meant that those individuals too would be likely to be able to perform the work of a High Court judge with aplomb.

When we spoke to solicitors, to lawyers in government service and to serving or retired County Court judges, we were frequently told that the myth that senior barristers must be better equipped for a High Court appointment was just that – a myth. One solicitor told us that:

> . . . the skillset for the High Court puts an emphasis on advocacy . . . court room experience . . . when in truth the ability to analyse, identify the issues, be methodical and logical and to communicate well and to manage . . . all those things solicitors do bloody well . . . but the job is described in the context of being on your feet, in court.

This connected to a wider view among some non-barrister interviewees that, in the eyes of the High Court judges, their role is of less importance than that of barristers: that solicitors are in effect 'second-class lawyers' and that this is reflected in a range of petty restrictions such as solicitors not being able to even announce in court that a case has been settled, to wear any kind of robe in court or to be addressed as 'my *learned* friend'. This view seems to cohere with that of senior judicial figures in England and Wales, as enunciated by Lord Neuberger when he argued that 'there is still an in-built assumption that it will tend to be a barrister, not a solicitor, who becomes a High Court judge . . . people on the whole do not think of solicitors – and solicitors do not think of themselves – as becoming High Court judges'.[41]

41 *Annual Oral Evidence to the House of Lords Select Committee on the Constitution taken on 29 March 2017* (Session 2016–2017) Q10 (Lord Neuberger of Abbotsbury) <www.parliament.uk/documents/lords-committees/constitution/Annual-evidence-2016-17/ucCC290317SupremeCourt.pdf>.

Naturally, such attitudes are somewhat resented. Solicitors and County Court judges were of the view that applicants from their own professions were often as likely to have the same or better skillsets compared with those of senior barristers. They reminded us that such individuals in their professions will have had to take many important decisions in the course of their careers, that solicitors are likely to be better than barristers at interacting with clients and that County Court judges will by definition be more knowledgeable about and practised in the art of 'court craft' or 'judge craft', that is, the skills required to keep cases running efficiently through the court while ensuring that all points needing to be addressed in each case are dealt with in appropriate detail both during the trial and in any resulting judgment.

The report by Turenne and Bell[42] did not specifically address the traditional preference given to senior counsel in competitions for High Court appointments. Nevertheless, many of the barriers to appointment which are discussed there apply just as much if not more so to non-traditional recruitment pools. The quantity, complexity and unremittingness of High Court work are as unlikely to be attractive to some non-traditional groups, such as senior solicitors, as they are to experienced barristers. The inflexibility of working hours and practices may be even less appealing to solicitors and lawyers in government service, who may have become used to benefiting from more sympathetic approaches to their caring responsibilities, their health needs and their work-life balance. One of our barrister interviewees commented that he or she would be reluctant to apply for the High Court partly because he or she periodically suffers from depression and did not think that reasonable adjustments would be made to take account of this disability.

The suitability of County Court judge applicants

Major advantages associated with encouraging applications to the High Court from the County Court bench are that a reduction in salary would be unlikely to play any part in the decision to apply for the job, the applicants would already be familiar with the trappings of judicial office (including the security implications, which are much more significant in Northern Ireland than in England and Wales) and well aware of the responsibilities and 'sacrifices' connected to holding such office. Many of them would be accustomed to the practice of writing court judgments and to the risk of adverse publicity arising out of their handling of a case, and, if they have been appointed to the County Court bench since the establishment of NIJAC in 2005, they would have undergone at least one previous application process and not be as daunted by it as some senior counsel who may have subjected themselves to something similar only when applying for silk. In addition, some County Court judges were previously solicitors, not barristers, so their appointment to the High Court would further assist in embedding diversity at that level. As of June 2019, there were 18 salaried County Court judges in Northern Ireland and 16 fee-paid deputy County Court judges. This constitutes a sizeable pool of potential applicants for High Court positions.

Unfortunately, there is a prevalent view among some barristers and a small number of serving High Court judges that, to be blunt about it, County Court judges would not be intellectually up to the task of serving in the High Court. That is certainly not the official line, as is evidenced by the fact that during the last year or so the Lord Chief Justice has exercised his power to invite a number of County Court judges to spend a number of

42 See n 3 above.

weeks within the Royal Courts of Justice acting as deputy High Court judges.[43] It seems that a prime motivation for this initiative may have been to meet the business needs of the High Court, particularly given that a vacancy in the High Court has existed for some time. However, it has provided a cadre of County Court judges with the experience of undertaking High Court work (even if, while serving in that Court, they were not paid at the same rate as High Court judges).

Given the shortness of their time in the High Court, it has not always been possible, we were told, to expose the deputy High Court judges to 'big' (i.e. complex or high-profile) cases. This suggests that perhaps this practice is not fulfilling its potential to develop an applicant pool among County Court judges. As one High Court judge put it, these deputy High Court judges 'should be given some of the more heavyweight work to demonstrate that they have the wherewithal to do it . . . [I]t would be helpful for the system to give them cases . . . where they can demonstrate their ability'. Nevertheless, those who have undertaken the role do seem to have welcomed the opportunity to do so and have seen it as a way of testing whether the role is attractive. Of course, if the current experiment were to be developed more systematically as a *recruitment* initiative, rather than as a way of managing business, it would be important that the system offered opportunities on an equal basis. Whether this should involve a role for NIJAC is open to debate, but there is a strong case for saying that it should.

Notwithstanding this recent experiment, there appears to be a view that County Court judges would be unlikely to be able to cope with the pressures of High Court work. It was suggested to us by several interviewees (including one High Court judge) that, if County Court judges were to be appointed to the High Court, they may not be adroit enough at handling fleet-footed senior counsel or at disaggregating the separate legal issues entwined in complex litigation. One senior counsel told us that even a County Court judge with 25 years' experience would not be adequately equipped to deal with what goes on in the High Court. On the other hand, one County Court judge described as 'absolute tripe' the suggestion that senior counsel would inevitably be better than County Court judges at dealing with knotty legal issues. An interviewee with many years of experience in more than one tier of the judiciary did not know why a County Court judge would be more likely to be 'bested' in a High Court hearing by a leading Queen's Counsel than a High Court judge would be, while another suggested that the idea of a County Court judge having the 'wool pulled over their eyes' by an experienced silk was 'a nonsense'.

Negative views about the County Court judiciary were expressed despite the fact that several County Court judges have many years' experience of running Crown Court trials (usually with a jury), including trials where the charges have included murder or serious sexual offences. We were reminded that summing up to a jury can be an 'extremely difficult' task. Criminal law now forms a small part of the workload of the High Court, so perhaps substantial experience in that field is not given as much weight in the selection process as experience in civil or family work is given. In addition, County Court judges hear all the extradition applications in Northern Ireland, which can raise quite complex points of law. They also regularly deal with difficult family law cases in the Family Care Centres, and, in that context, they may have to manage many litigants in person. Difficult

43 Section 7(2) of the Judicature (NI) Act 1978 reads: 'A county court judge shall, if requested to do so by the Lord Chief Justice, sit and act as a judge of the High Court.' Compare the power conferred by s 7(3): see n 51 below.

points can arise in civil litigation at the County Court level too, as exemplified all too well when the 'gay cake' case was first litigated in that forum.[44]

A County Court judge was firmly of the view that, because of the consistent failure of County Court judges to be appointed to the High Court in the last decade, there was now a chill factor amongst serving County Court judges which enhanced their reluctance to apply to the High Court. That interviewee knew of one 'very good' County Court judge who applied but did not get to the final assessment stage. A senior counsel – who had already tried to be appointed to the County Court but unsuccessfully – argued strongly that several of the existing County Court judges were easily as good lawyers as some senior counsel.

In England and Wales, Lord Chief Justice Burnett has let it be known that '[p]roper appraisals, consistent with judicial independence, are being rolled out across the judiciary within the resources available to us'.[45] We are not aware that the same is happening in Northern Ireland. If it were, then County Court judges who are led to believe from their appraisal that they are performing their role particularly well might be more incentivised to apply for a High Court position in due course. Lord Burnett also said:

> Career progression within the judiciary is an important factor in increasing diversity at the higher levels. The appointments process should recognise that in the ranks of both salaried and fee-paid judges are many individuals who have not been visible in the higher ranks of the professions, often because they have chosen a career path in the judiciary which is more readily compatible with family life, but who have demonstrated judicial ability.[46]

As examples of the kind of steps being taken in England and Wales to encourage applications for judicial posts Lord Burnett referred to specific pre-application seminars, a Pre-Application (Online) Judicial Education Programme, a Judicial Work Shadowing Scheme, a formal judicial-mentoring scheme, and a scheme enabling individuals to apply for a fixed-term appointment as a temporary High Court judge.[47] Regarding the last of these initiatives, he pointed out that by 2018 a total of 73 deputy High Court judges had been appointed and that nine of these had subsequently been appointed as permanent High Court judges.[48]

The ban on former judges returning to practice

Applicants for High Court appointments are referred to a document on NIJAC's website called 'Terms and Conditions'. This makes it clear, in paragraph 2.1, that 'The Lord Chancellor regards appointment to the Bench as being for life. Any offer of appointment is therefore made on the understanding that the appointee will not return to practice.' This prohibition on future practice affects all categories of applicant equally, but it is perhaps even more off-putting for younger applicants since resigning from their judicial post after a relatively short time on the bench will entail a greater sacrifice in career terms than it would for an older person. Former High Court judges can become chairs of

44 *Lee v Ashers Baking Co Ltd* [2015] NICty 2, heard by District Judge Brownlie while sitting as a deputy County Court judge; the case later went to the Court of Appeal ([2016] NICA 23 and 55) and then to the UK Supreme Court ([2018] UKSC 49, [2018] 3 WLR 1294).

45 The Treasurer's Lecture (n 10) 23.

46 Ibid 26.

47 Ibid 27–29.

48 Ibid 29.

inquiries (Judges Hart and Coghlin are recent examples), arbitrators, mediators or in-house lawyers, but otherwise they could find it difficult to make a living out of their legal knowledge.

This potential barrier to judicial appointment was not mentioned to us by any interviewee, perhaps because it affects persons who are thinking of applying to lower-tier judicial appointments more so than it does prospective applicants to the High Court. However, Turenne and Bell observe that a number of the people who responded to their question on the issue thought that the prohibition on future practice was 'an unnecessary limitation'.[49] They also note that the House of Lords Select Committee on the Constitution had received representations from both the Bar and solicitors to the effect that the ban on returning to practice was a disincentive to potential applicants for judicial posts.

Part-time and 'flexitime' High Court judges

This context provides a further justification for the appointment by NIJAC of part-time High Court judges. While previous research on judicial appointments in Northern Ireland suggests that recruitment to judicial posts on a part-time basis may not be appropriate in all circumstances,[50] there was some support amongst our respondents for part-time working arrangements at the High Court level. Facilitating part-time working can, of course, be of particular benefit to people with dependents, such as those with caring responsibilities. There are already 16 fee-paid part-time County Court judges and 23 fee-paid part-time District Judges (Magistrates' Courts), all of whom can be called 'deputy' judges.[51] They are appointed for five-year terms, although these are renewable provided the appointees retire at the age of 70. We understand that a few other salaried judges, such as Employment Judges who sit in industrial tribunals, have reduced the number of hours they work and so are now part-time.

It is not currently possible for an advertisement for a permanent High Court post in Northern Ireland to guarantee that the person appointed will be able to work on a part-time basis. Legislation stipulates a maximum number of puisne High Court *judges* (10), not a maximum number of full-time equivalent High Court *posts*.[52] Additional revenue would be required from HM Treasury if more than 10 individuals were to serve as permanent full-time or part-time High Court judges.

But as mentioned above,[53] section 7(2) of the Judicature (NI) Act 1978 allows the Lord Chief Justice to request a County Court judge to sit and act as a judge of the High Court, and in exercise of that power several County Court judges have recently been invited to serve in that capacity for a few weeks or months. In addition, just after the research discussed in this paper was completed, NIJAC has for the first time approved a method for recruiting up to 10 temporary High Court judges under section 7(3) of the Judicature (NI) Act 1978. This permits NIJAC to appoint persons 'to sit and act as a judge of the High Court as a temporary measure in order to facilitate the disposal of business in the High Court or the Crown Court'. Although the outcome of the competition for

49 Above, n 4 para 101.

50 Above, n 5 at 6.

51 The number of judges was provided by NIJAC, accurate as of 18 April 2019. The relevant legislation does not impose any numbers limits. Deputy County Court judges are appointed by NIJAC under the County Courts Act (NI) 1959, s 107 (as amended), while Deputy District Judges (magistrates' courts) are appointed by NIJAC under the Magistrates' Courts Act (NI) 1964, s 10 (as amended).

52 Judicature (NI) Act 1978, s 2(1), as amended by the Maximum Number of Judges (NI) Order 2004 (SI 2004/1985).

53 Above pp 494–495.

these temporary posts is not yet clear, the development is one which should be welcomed in principle. We understand from informal discussions with two members of the Bar that the level of interest in the appointments appears to have been high. It is anticipated that in the early part of 2020 a number of temporary High Court judges from varied backgrounds in legal practice will be sworn in. They will initially be appointed for three years, but that period will be renewable, as it is for deputy County Court judges and deputy District Judges (magistrates' courts). When future opportunities arise to compete for positions as permanent High Court judges (there are currently two vacancies within the statutory complement of 10), the temporary judges, as well as some individuals who were unsuccessful in the competition for the temporary posts, may well apply.

We discussed with several interviewees whether they thought it would be feasible for a practising barrister or solicitor to try out the role of High Court judge while still maintaining their practice. The consensus (but not the unanimous view) was that it would *not* be feasible, due not only to the likely clash of diary commitments but also to the potential conflict of interests that might arise in a small jurisdiction like Northern Ireland. But one QC did serve as a deputy High Court judge in the 1990s while still maintaining his practice, and some of the 16 persons who have been appointed as deputy County Court judges continue to work as practising barristers or solicitors. If such 'double-jobbing' is possible at that level, it should also be possible at the High Court level. Such positions may be attractive to practitioners who are nearing the end of their career in practice but who wish to experience one new challenge before retirement. The appointments should help to reduce the workload falling on permanent High Court judges, thereby in due course making their job more attractive to future applicants. Deputy High Court judges who are otherwise practising lawyers (and/or academics)[54] are common in England and Wales, but we gather that it is easier in that larger jurisdiction for practising lawyers to undertake specialised High Court work on a trial basis without this making it very difficult for them to carry on their own private practice.

The UK government's Advisory Panel on Judicial Diversity produced a report in 2010 which stated that 'it should be assumed that all posts are capable of being delivered through some form of flexible working arrangements, with exceptions needing justifying'.[55] Of course, quite apart from working on a temporary or part-time basis, all employees, including judges, have the right to request flexible working hours, although no right to demand them.[56] The request can be made only once a year, but the employer may refuse it only if he or she considers that one or more of a number of stipulated grounds applies, such as that the request would lead to additional costs or would detrimentally affect the ability to meet customer demand. Changes that might allow flexible working were found to have a particular appeal to several of our female interviewees. We understand that the Lord Chief Justice of Northern Ireland is willing to consider requests from High Court judges for some degree of flexibility over holiday periods and days off, but we are not aware that any formal flexitime arrangement has been agreed with any of the current High Court judges.

54 On 24 July 2019 it was announced that Professor Andrew Burrows, Professor of the Law of England at the University of Oxford and a deputy High Court judge, was to be appointed to the UK Supreme Court from June 2020.

55 *Report of the Advisory Panel on Judicial Diversity* (2010), Recommendation 51.

56 See the Employment Rights (NI) Order 1996, Arts 112F–112I, inserted by the Employment (NI) Oder 2002, Art 15, and the Flexible Working Regs (NI) 2015 (SR 105).

Judging as a career

One of the clearest messages we received during the course of our research was that it was remarkable that no County Court judge had made it to the High Court in Northern Ireland since the new NIJAC-run selection system was put in place in 2005. It was suggested to us that the chances of the equivalent level of judges (Circuit judges) being appointed to the High Court in England and Wales are roughly the same (at about one in ten), but that of itself is hardly a justification for the phenomenon in Northern Ireland. It is worth noting that, in the Republic of Ireland, new developments suggest that serving judges in the superior courts will not have to apply at all through the Judicial Appointments Commission for promotion to the High Court, Court of Appeal or Supreme Court, following a significant concession by the Irish government to opponents of the Judicial Appointments Commission Bill that is currently passing through the Oireachtas.[57] Instead, serving judges will automatically be considered for promotion.

To the informed outsider – someone who is not familiar with the prevailing culture within the various branches of the legal professions – the fact that 'junior' judges do not regularly get appointed as 'senior' judges must seem bizarre. Most people would assume that the persons most likely to make good High Court judges would be those who have already demonstrated their judging abilities in a lower court. As one County Court judge told us, 'the Bar should never be seen as the sole pool or even necessarily the best pool for higher judicial office . . . the skills don't necessarily transfer over.'

Considering the relative skillsets of experienced senior counsel and experienced County Court judges (several of whom would have been senior counsel before being appointed to the County Court), as well as the conversations we have had with several interviewees who have experience of judicial work in the County Court, it may well be that something in the selection process for High Court posts in Northern Ireland is working to the disadvantage of applicants from the County Court. We deduce that the judicial experience already gained by such applicants is perhaps not being given enough weight in the process. While it may be understandable that judges on the Selection Committee who may have come to the High Court through the traditional route (i.e. directly from the ranks of Queen's Counsel) might be prone to unconsciously favour applicants who have a similar background to themselves, it is surprising that lay members of the Selection Committee do not set greater store by an applicant's existing judicial experience when considering their suitability for the High Court.

If in future a common route to the High Court were to consist of many years' experience in the practice of law, in whatever capacity, coupled with several years' experience as a judge at a lower tier in the court system, this would undoubtedly be unpopular within the Bar, especially the senior Bar, but it may well lead to better qualified applicants competing for appointments. Senior barristers have the choice – when they have already been earning high sums for some time – to attempt to divert into the lower tier of the judicial profession if they so wish. The choice is not very different from that facing longstanding junior counsel when confronted with the prospect of applying to take silk: presumably they wish to enjoy the esteem which goes with the move, but they are unsure if they will be able to make a success of the new role. Lawyers who opt to join the lower tier of the judiciary should not then have to conclude that their chances of progressing further up the judicial ladder are no better than one in ten.

We were also made aware by several interviewees, especially those from government legal services (e.g. the Public Prosecution Service) and those attending a meeting of

57 *Irish Times* (Dublin, 9 May 2019).

'Presiders',[58] that there seems to be little tradition in Northern Ireland of persons employed as judges in tribunals or as District Judges being promoted to a higher tier of the judiciary.[59] NIJAC informed us that three individuals who served as fee-paid legal members of tribunals and then as salaried chairs of tribunals subsequently applied successfully to become County Court judges, but the Commission confirmed that to date it has never selected a District Judge (magistrates' courts) to be promoted to the level of full-time County Court judge. That being said, the four District Judges who operate in the civil courts have all been appointed deputy County Court judges as well, and one of them recently became the first such District Judge since NIJAC's establishment in 2005 to be promoted to the position of County Court judge.

We accept that, while it is important that judge-applicants have a proper opportunity to reveal to the Selection Committee the skills they have already acquired while serving as judges, the competition for High Court appointments should not be changed so as to be skewed in favour of judge-applicants. Rather, the skills in question need to be ones which lawyers who have not yet served as judges could still demonstrate based on their legal experience to date. We suggest that these skills could be: (a) an ability to manage a number of cases that need to be processed more or less at the same time; (b) an ability to manage particular cases effectively and efficiently, including an ability to interact successfully with legal practitioners involved in the same case; (c) an ability to make decisions; and (d) an ability to explain decisions in a convincing manner both orally at the time and in writing later.

6 Conclusion

The research which has been outlined in this paper bears out, and expands upon, the findings of research conducted in England and Wales. While gender was not a particular focus of this study as it had been in earlier work,[60] it is still worth considering how many of the barriers may apply to women to a greater degree than they do to men.

Chief among the issues identified in other studies was the negative impact which the current provisions for judicial pay and pensions have had on recruitment to the High Court. This obtains in relation to appointments to High Court positions in Northern Ireland too. Changes require to be made to those provisions. But the research also confirms that there are several other barriers to recruitment which are not related to personal financial circumstances. The legal professions in Northern Ireland, as elsewhere, have changed: becoming a judge is no longer necessarily seen as the pinnacle of a successful career (generally at the Bar). Legal practice in both branches of the profession can be very rewarding without inevitably requiring the sort of commitment of time and energy that is expected of a judge as he or she reaches the final stages of their career. Also, the nature of the work on the bench, sometimes involving historical sex abuse or troubles-related legacy issues, as well as an increase in litigants in person, may for some be less appealing than continuing in specialist practice. Modern working conditions within both branches of the legal profession now involve a degree of flexibility that is not always fully reproduced within the 'back corridor culture' of the High Court, with its limited legal research support structures. Furthermore, the absence of the traditional 'tap on the shoulder' means that there is not always the same pressure as in the past to take one's turn to give something back to the system. There is also a fear of failure and, perhaps most

58 That is, the people who preside over different tiers of the judiciary, whether in courts or in tribunals.

59 In civil law countries, of course, such promotion is normal because they prefer to have a 'career judiciary'.

60 See above n 4.

significantly, a continuing belief that the main pool for successful applicants will continue to be the senior Bar. This undoubtedly creates a 'chill factor' within those from the County Court or government service who might otherwise be less susceptible to many of the factors that the senior Bar find off-putting.

For some these issues are not enough to deter them from a career in the High Court: the challenge of new work, the sense of public duty and a change in the nature of the pressure of workload were cited as appealing by most of the serving High Court judges we interviewed. However, this appears to be true for fewer potential applicants than in the past, and new pools of talent from existing judicial offices and public service are not replacing the shortfall from the traditional quarters.

While the process of appointment is seen by some potential applicants as daunting, it is generally not perceived as an insurmountable obstacle, although perhaps it should become more widely known that multiple applications may be required before success ensues. Several further steps could be taken to make the High Court recruitment process speedier, fairer and more attractive. Perhaps the most important of these is the development of the temporary position of deputy High Court judge. Some of the people serving in that role could in turn constitute a well-qualified pool of potential applicants for permanent High Court positions. Furthermore, and perhaps this is one of the more important findings, if some of those permanent positions could be available on a part-time or flexitime basis the chances of recruiting a more diverse but still very talented group of High Court judges would be enhanced still further.

Perhaps our most significant finding relates to the issue of the absence of a judicial career path that might resolve the problem of insufficient traditional applicants from the senior Bar. Traditional understandings about progression from the senior Bar to the High Court are no longer delivering the volume of suitable applicants, but existing 'junior' judges from lower tiers are not filling the gap as might be expected in any other field where career progression is often the norm. While not every County Court judge will be suitable for the High Court, there seems to be a strong sense among potential applicants, not only from the County Court but also to an extent from the solicitors' profession and lawyers in public service, that their skills are not as highly valued and that applications from these quarters are less likely to be successful. This means that such applicants are less willing to come forward and, while NIJAC can only consider for appointment those who actually apply, there does appear to be a blockage in the system that requires action to clear. In circumstances where financial, structural and cultural changes to the role both of senior lawyers and of High Court judges limit the capacity of NIJAC as the appointing body to engender sufficient applications of quality from traditional pools, it is vital that the net is widened – by whatever means possible, including most notably the actual appointment of one or more good applicants from the County Court – to ensure that the system is open to catching the most meritorious applicants from wherever they may be currently working.

NILQ 70(4): 505–510

The Irish Supreme Court judgment in *Nano Nagle School v Marie Daly*: a saga of litigation

Dr Shivaun Quinlivan

and

Dr Charles O'Mahony

National University of Ireland Galway

1 Introduction

The decision of the Supreme Court in the case of *Nano Nagle School v Daly* is of significant importance for employees and employers, as well as for the Labour Court.[1] A key implication of the Supreme Court judgment is that reasonable accommodation can entail the redistribution of duties, subject, of course, to the statutory limitation that this accommodation does not impose a disproportionate burden on the employer. Flowing from this, the Supreme Court held that employers are under an obligation to consider the impact of a proposed redistribution of duties before they can refuse to provide reasonable accommodation. Moreover, the wise employer would meaningfully consult with the employee concerned in the discharge of their duty of reasonable accommodation.

A further key element of this decision, and a marked departure by the Supreme Court from its relatively recent decisions, is its engagement with the relevant EU law.[2] In this judgment the Supreme Court not only referenced the Court of Justice of the European Union's analysis of the Framework Employment Directive[3] but also accepted that the Directive must be interpreted in light of the UN Convention on the Rights of Persons with Disabilities (CRPD).[4] This opens up the potential for the CRPD to impact upon the interpretation of Irish employment equality law.

2 Background to this case

This case relates to Marie Daly, a special needs assistant (SNA) who worked for the Nano Nagle School in County Kerry. The school operates under the aegis of the National Council for Special Education (NCSE), which is a state agency and a funding authority. The school has operated since 1998 for children with learning and/or physical disabilities.

1 *Nano Nagle v Daly* [2019] IESC 63.

2 For a discussion of this issue see Mel Cousins, 'Education and the Equal Status Acts: Stokes v Christian Brothers High School Clonmel' (2015) 38 Dublin University Law Journal 157; Shivaun Quinlivan, 'Stokes v Christian Brothers High School: An Exercise in Splendid Isolationism?' (2015) 18(2) International Journal of European Law 81.

3 Council Directive 2000/78/EC of 27 November 2000 establishing a general framework for equal treatment in employment and occupation (hereafter Framework Employment Directive).

4 UN General Assembly, Convention on the Rights of Persons with Disabilities: Resolution/Adopted by the General Assembly, 24 January 2007, A/RES/61/106.

In 2010 the appellant Marie Daly was involved in an accident while on holiday, as a result of which she was paralysed from the waist down. Following an intensive period of rehabilitation, in 2011 she sought to return to work. In a subsequent review of expert reports, the school board concluded that she did not have the capacity to undertake the full set of duties associated with her post as an SNA and that she would not be able to do so in the future. The school declined to allow her return to work. The appellant made a complaint to the Equality Tribunal (now subsumed within the Workplace Relations Commission (WRC)) on the basis that the school had failed to provide appropriate measures to accommodate her, as a person with a disability, to return to work contrary to section 16 of the Employment Equality Acts 1998–2015.

The case was described by Ryan P in the Court of Appeal as a 'saga of litigation'[5] and by MacMenamin J in the Supreme Court as 'overlong litigation'.[6] As will be discussed below, the Supreme Court has now remitted the case back to the Labour Court to resolve. The case had been considered first by the Equality Tribunal (now WRC); second by the Labour Court on appeal; third by the High Court, by way of an appeal on a point of law; fourth by way of appeal to the Court of Appeal; and most recently on appeal to the Supreme Court.

One of the core legal issues that came before the different tribunals/courts in the *Nano Nagle* case related to the interpretation of section 16 of the Employment Equality Acts, which addresses the duty to provide 'appropriate measures' more commonly described as the duty to provide reasonable accommodation. Specifically, the question was whether the school was in breach of the obligations imposed by that section in dismissing the appellant from her employment. The Equality Tribunal delivered its decision in 2013 and determined the case in favour of the employer. Marie Daly appealed the decision to the Labour Court, which delivered its determination in 2014. The Labour Court determined the case in favour of Marie Daly and awarded her €40,000 by way of compensation. It held that the school 'had a duty to fully consider the viability of a reorganisation of work and a redistribution of tasks among all of the SNAs so as to relieve [Ms Daly] of those duties that she was unable to perform'.[7] This determination was then appealed to the High Court on a point of law.

In the High Court, Noonan J upheld the determination of the Labour Court. He held that the school was in breach of section 16 of the Employment Equality Acts on the basis that the employer had failed to consider a redistribution of tasks. This decision was further appealed to the Court of Appeal. Ryan P, Finlay Geoghegan J and Birmingham J delivered their judgments in 2018 allowing the appeal and setting aside the determination of the Labour Court.

The judgment by the Court of Appeal was criticised as not providing full or appropriate protection for employees with disabilities.[8] In particular, the court's interpretation of section 16 of the Employment Equality Acts did not meet the minimum standard required by the Framework Employment Directive, which requires reasonable

5 *Nano Nagle School v Daly* [2018] IECA 1.

6 *Nano Nagle v Daly* [2019] IESC 63, para 111.

7 *A Worker (Complainant) v A School (Respondent)* [2014] 25 ELR 307, 342–343.

8 Claire Bruton and Katherine McVeigh, 'Effects of the Judgment of the Court of Appeal in Nano Nagle v Daly on the Duty to Provide Reasonable Accommodation' (2018) 15(2) International European Law Journal 36–47.

accommodation to be interpreted broadly.[9] A related concern was the Court of Appeal's distinction between two terms used in that section: 'tasks' and 'duties', specifically equating 'duties' with 'essential functions'.[10] The Court of Appeal held that an employer was only required to consider the redistribution of 'tasks' and was not required to consider the redistribution of 'duties'. Therefore, the distinction between the terms 'tasks' and 'duties' becomes important as it 'marks the line between protection and no protection for employees with a disability'.[11]

Unfortunately, the Court of Appeal failed to provide adequate clarification on how to differentiate between those terms. In this, it adopted an unduly restrictive and narrow interpretation of the reasonable accommodation duty. In effect as Bruton and McVeigh stated, the Court of Appeal focused on the perceived 'incompetency and incapability of a person with a disability . . . as opposed to the abilities of employees with disabilities with the provision of reasonable accommodation'.[12] The judgment of the Supreme Court is important in rejecting the Court of Appeal's narrow interpretation of section 16(1), which effectively restricted the duty to provide reasonable accommodation and ran counter to the purpose of introducing this legal duty into Irish employment equality law.

3 The Supreme Court decision

3.1 The test for reasonable accommodation

The duty to provide reasonable accommodation can be broken down into two constituent parts.[13] The first part imposes a positive legal duty to provide a reasonable accommodation where required. The second part of this duty ensures that those required accommodations do not impose a disproportionate burden on the employer. With regards to the first part of this test, the Supreme Court determined that this was a question of fact and one that was for the Labour Court to determine. On the second part, of what would give rise to a disproportionate burden, the Supreme Court provided useful guidance.

In determining what is a disproportionate burden, the Supreme Court noted that the duty to accommodate was neither 'infinite, [nor] at large'.[14] In determining that there was no real distinction to be made between 'tasks' and 'duties', there was 'no reason, in principle, why certain work 'duties' cannot be removed or 'stripped out'.[15] That stated, the subtraction of certain duties cannot result in removing all duties, because that would

9 *HK Danmark, acting on behalf of Jette Ring v Dansk Almennyttigt Boligselskab and HK Danmark, acting on behalf of Lone Skouboe Werge v Dansk Arbejdsgiverforening, acting on behalf of Pro Display A/S (Ring and Skouboe Werge)* (2013) Joined Cases C-335/11 and C-337/11.

10 Council Directive 2000/78/EC of 27 November 2000 establishing a general framework for equal treatment in employment and occupation establishes that the duty to recruit, promote, maintain in employment or train an individual only applies to those willing, capable and competent to undertake the 'essential functions' of the post. This is without prejudice to the obligation to provide reasonable accommodation for people with disabilities.

11 Bruton and McVeigh (n 8).

12 Ibid.

13 See Shivaun Quinlivan, 'Reasonable Accommodation: An Integral Part of the Right to Education for Persons with Disabilities' in Gauthier De Beco, Shivaun Quinlivan and Janet E Lord (eds), *The Right to Inclusive Education in International Human Rights Law* (Cambridge University Press 2019) 169–189, 175; UN Committee on the Rights of Persons with Disabilities, *General Comment No 6: Equality and Non-discrimination* (2018) CRPD/C/GC/6, para 25.

14 *Nano Nagle v Daly* [2019] IESC 63, 89.

15 Ibid.

almost inevitably become a 'disproportionate burden'.[16] The test to determine the extent of the duty must be one of fact and is guided by the principles of 'reasonableness and proportionality'.[17]

3.2 The duty to consult

The issue of the duty to consult with an employee in determining whether or not a reasonable accommodation is possible was previously considered by Dunne J in the Circuit Court case of *Humphries v Westwood*.[18] In that case, she determined the nature and extent of the enquiry that an employer should undertake to determine whether an employee was fully capable of undertaking the duties of the post and what, if any, special treatment or facilities might be available by which the employee could become fully capable. Dunne J went on to hold that such 'enquiry could only be regarded as adequate if the employee concerned is allowed a full opportunity to participate at each level and is allowed to present relevant medical evidence and submissions'.[19] In this case it was clear that the school principal did not consult with Ms Daly in reaching the decision that the 'appellant could return to work only if she was able to perform all the duties of an SNA, with assistance or otherwise'.[20] The Labour Court held that the school had failed to comply with section 16(3) of the Act as it had failed to consult with the appellant and awarded her €40,000 in compensation for this failure. In contrast, the Court of Appeal rejected the notion that a failure to consult constituted a breach of the duty imposed.[21] The Supreme Court addressed whether dialogue with a person with a disability was a necessary element of the duty to provide reasonable accommodation. While accepting the wisdom of including the person with a disability in the process, the Supreme Court held that the 'absence of consultation cannot, in itself, constitute discrimination under s.8 of the Act'.[22]

3.3 Implications of the judgment for Employment Tribunals

The Supreme Court in its judgment made a number of statements that have implications for tribunals adjudicating on employment law. It is of note that the Supreme Court acknowledged that the courts exercise a measure of deference 'to an administrative tribunal acting within the scope of its duty'.[23] However, it also noted that, where there is a substantial failure by a tribunal to comply with its statutory duty, a court must intervene. The Supreme Court was critical of the Labour Court's handling of this case in that it had omitted any reference to expert evidence described as highly important and capable of having a direct bearing on the outcome of the case. The Supreme Court described this as a 'factual lacuna'.[24]

In order to address this lacuna the Supreme Court decided to remit the appeal back to the Labour Court, which is the body that has the mandate to evaluate evidence in accordance with the relevant legislation, and assess its significance. The rationale for this was that the issues around the omitted evidence could only be resolved by the Labour

16 Ibid.
17 Ibid.
18 [2004] 15 ELR 296.
19 Ibid 297.
20 *Nano Nagle v Daly* [2019] IESC 63, 58.
21 *Nano Nagle School v Daly* [2018] IECA 54.
22 *Nano Nagle v Daly* [2019] IESC 63, 105
23 Ibid 109.
24 Ibid 108.

Court itself. The Supreme Court also noted that it 'should not act as a surrogate Labour Court, which is charged with carrying out a statutory function'.[25] Given the complex and lengthy litigation preceding the judgment, the Supreme Court did not take the decision to remit the case back to the Labour Court lightly. Nevertheless, this saga of litigation continues.

3.4 EU LAW

A number of elements of the Supreme Court decision are to be welcomed. One particularly welcome development was the engagement by the Supreme Court with the provisions of both the Framework Employment Directive and the CRPD. In this respect the partial dissent of Charleton J is particularly welcome where he unequivocally accepts the impact of both the Framework Employment Directive and the CRPD, stating that the '1998 Act must be interpreted in the light of the Directive. In turn, the Directive requires to be seen in the light of the international obligations entered into by the European Union, the Convention.'[26]

The majority are perhaps more tentative on this point, MacMenamin J (O'Donnell, Dunne and O'Malley JJ concurring) stating early in their judgment: 'EU law is undoubtedly a useful point of reference. But whether it is even necessary to resort to EU law is a point to be determined.'[27] However, MacMenamin J did reference the CRPD and its impact on the Framework Employment Directive in his judgment. Ultimately, MacMenamin J decided that the case could be determined by reference to the wording of the relevant section alone, and there was no necessity to have 'resort to the judgment of the CJEU in *Ring*, or the Framework Directive'.[28] While deciding the issue of principle in this case on the basis of statutory interpretation he went on to state:

> This conclusion, based on the words of the section alone, as it happens, accords with any interpretation of the section by reference to the reasoning of the CJEU in *Ring*.[29]

The fact that the majority of the Supreme Court engaged, albeit tentatively, with Ireland's obligations under the Framework Employment Directive appears to be a significant change in approach since its decision in *Stokes v Christian Brothers High School Clonmel*.[30] In that case the Supreme Court addressed the provisions of the Equal Status Acts 2000–2015 as though only a matter of national statutory interpretation were involved, rather than giving effect to an EU obligation.

The Supreme Court appears to accept that both the CRPD and the Framework Employment Directive have relevance to this issue. It is of note that the UN CRPD Committee has clearly stated that a key element of the duty to provide reasonable accommodation includes 'dialogue with the person with a disability concerned'.[31] While accepting that dialogue may be the prudent approach for employers, the Supreme Court does not impose a duty to consult. In addition, the CRPD Committee has held that reasonable accommodation is a term of art, and the term 'reasonable' should not act as

25 Ibid 111.

26 *Nano Nagle v Daly* [2019] IESC 63, Charleton J dissenting at 3.

27 Ibid 16.

28 Ibid 91.

29 Ibid.

30 [2015] IESC 13.

31 UN Committee on the Rights of Persons with Disabilities (n 13) para 26(a).

32 Ibid para 25(a).

a 'distinct qualifier or modifier of the duty';[32] rather the Committee has held that the reasonableness of an accommodation relates to the 'relevance, appropriateness and effectiveness for the person with a disability'.[33] This is at odds with the Supreme Court's position which held that the test of disproportionate burden was one of 'reasonableness and proportionality'.[34] It is suggested by the authors that this case may not be the last word on this issue.

4 Conclusion

The Supreme Court judgment in this case is an important one. First, it overturns the troubling judgment of the Court of Appeal, which adopted an overly restrictive and narrow interpretation of reasonable accommodation. The Supreme Court held that reasonable accommodation can involve a redistribution of any task or duty in a particular job, provided it does not impose a disproportionate burden on the employer. Second, the Supreme Court addressed whether dialogue with an employee with a disability was a necessary element of the duty to provide reasonable accommodation. The Supreme Court noted that the prudent employer would consult. Third, the Supreme Court's serious engagement with both the Framework Employment Directive and CRPD highlights the emerging importance of these sources of law in determining questions of domestic employment equality law.

33 Ibid.

34 *Nano Nagle v Daly* [2019] IESC 63, 89.

NILQ 70(4): 511–521

Liability threshold for damages in public procurement: the EFTA Court's *Fosen-Linjen* saga

ALBERT SANCHEZ-GRAELLS*

University of Bristol

Abstract

In this case comment, I explore the two EFTA Court Judgments in the Fosen-Linjen *saga and their opposing views on the interaction between EU/EEA rules on procurement remedies and the more general principle of state liability for breaches of EU/EEA law. I review the case law of the Court of Justice of the European Union and, in particular, the perceived inconsistencies between the two 2010 judgments in* Strabag *and* Spijker*, which featured very prominently in the legal arguments submitted to the EFTA Court in both* Fosen-Linjen *cases. I also use the benchmark of the UK Supreme Court's* Nuclear Decommissioning Authority *judgment to support the view that* Spijker *reflects the correct understanding of EU/EEA law and that there should be no further debate about it. I submit that the Court of Justice of the European Union would be well-advised to (re)confirm the position enshrined in* Spijker *at the earliest opportunity, to avoid any perpetuation of this debate in the context of EU/EEA public procurement law.*

Keywords: public procurement; damages; liability threshold; *Fosen-Linjen*; *Nuclear Decommissioning Authority*; EFTA Court; CJEU; UK Supreme Court; *Spijker*; *Strabag*

Introduction

On 1 August 2019, the European Free Trade Association's Court (EFTA Court) confirmed in *Fosen-Linjen II*[1] that the threshold for damages claims for breaches of EU/European Economic Area (EEA) public procurement law is that of a 'sufficiently

* Professor of Economic Law, University of Bristol Law School. This case note builds on the previous discussion in the paper 'You Can't Be Serious: Critical Reflections on the Liability Threshold for Damages Claims for Breach of EU Public Procurement Law after the EFTA Court's Fosen-Linjen Opinion' (2018) 1(1) Nordic Journal of European Law 1-23, which discussed the *Fosen-Linjen I* judgment in detail. I am grateful to Prof Roberto Caranta, Dr Roxana Vornicu and the contributors to a special issue of the European Procurement and Public Private Partnership Law Review for further discussions on the *Fosen-Linjen* saga. A complementary discussion of the saga from the perspective of the minimum harmonisation carried out by the Remedies Directive can be found there. The standard disclaimer applies. Further comments welcome: a.sanchez-graells@bristol.ac.uk.

1 Judgment of 1 August 2019 in *Fosen-Linjen AS, supported by Næringslivets Hovedorganisasjon (NHO) v AtB AS* (E-7/18) ('*Fosen-Linjen II*') <https://eftacourt.int/download/7-18-judgment/?wpdmdl=6235>.

serious breach' of the applicable rules and that the Remedies Directive[2] 'does not require that any breach of the rules governing public procurement in itself is sufficient to award damages'.[3] This judgment resulted from a request for an advisory opinion from the Supreme Court of Norway, which harboured doubts as to the adequacy of the earlier EFTA Court Judgment of 31 October 2017 in *Fosen-Linjen I*,[4] where the Court had taken the diametrically opposed view that a 'simple breach of public procurement law is in itself sufficient to trigger the liability of the contracting authority'.[5] This remarkable U-turn by the EFTA Court has raised significant controversy, not least because both advisory opinions were issued in the context of the same underlying litigation.

This case note considers the systematic interpretation issues that underpin the *Fosen-Linjen* saga and, in particular, the contested relationship between the system of remedies for the enforcement of EU/EEA public procurement law and the more general principle of state liability for breaches of EU/EEA law (i.e. the so-called *Francovich* doctrine).[6] Such interpretive issues stem from the perceived contradiction between two 2010 judgments of the Court of Justice of the European Union (CJEU) in *Strabag*[7] and *Spijker*.[8] Remarkably, this had already been addressed in 2017 by the UK Supreme Court in *Nuclear Decommissioning Authority v EnergySolutions*,[9] which reached the same position as *Fosen-Linjen II*.

1 The facts

In *Fosen-Linjen*, the contracting authority had tendered a contract for the provision of ferry services.[10] The award criteria were 'price' (50 per cent), 'environment' (25 per cent) and 'quality' (25 per cent).[11] Remarkably, the evaluation of the award criterion 'environment' was based on the tenderers' specification of fuel oil consumption for the offered ferries, but tenderers were not required to demonstrate how the fuel oil consumption value was calculated or to state the assumptions upon which the calculations were based.[12]

After being judicially challenged on this approach, the contracting authority eventually acknowledged that it had failed to establish a reasonable basis for evaluation and that it had committed an error by not verifying the reasonableness of the tenders, which could be in violation of the applicable EU/EEA procurement rules.[13] The contracting authority cancelled the tender and entered into an emergency contract for the short-term provision of ferry services, with a view to retendering the longer-term contract. The

2 Directive 89/665/EEC on the coordination of the laws, regulations and administrative provisions relating to the application of review procedures to the award of public supply and public works contracts [1989] OJ L395/33, as amended by Directive 92/50/EEC [1992] OJ L 209/1, and by Directive 2007/66/EC [2007] L 335/31 (hereinafter, the 'Remedies Directive').

3 *Fosen-Linjen II* (n 1) para 121.

4 Judgment of 31 October 2017 in *Fosen-Linjen AS v AtB AS* (E-16/16) ('*Fosen-Linjen I*') <https://eftacourt.int/wp-content/uploads/2019/01/16_16_Judgment_EN.pdf>.

5 Ibid para 82.

6 Following the judgments of 19 November 1991, *Francovich and Others*, C-6/90 and C-9/90, EU:C:1991:428, and of 5 March 1996, *Brasserie du Pêcheur and Factortame*, C-46/93 and C-48/93, EU:C:1996:79.

7 Judgment of 30 September 2010, *Strabag and Others*, C-314/09, EU:C:2010:567.

8 Judgment of 9 December 2010, *Combinatie Spijker Infrabouw-De Jonge Konstruktie and Others*, C-568/08, EU:C:2010:751.

9 *Nuclear Decommissioning Authority v EnergySolutions EU Ltd (now ATK Energy Ltd)* [2017] UKSC 34.

10 *Fosen-Linjen I* (n 4) para 19.

11 Ibid para 23.

12 Ibid para 25.

13 As interpreted in the Judgment of 4 December 2003, *EVN and Wiestrom*, C-448/01, EU:C:2003:651.

disappointed tenderer Fosen-Linjen did not submit a tender in the new procedure.[14] Instead, *Fosen-Linjen* sued the contracting authority in damages for positive contract interest (loss of profit – *lucrum cessans*) or, in the alternative, for negative contract interest (costs of bidding – *damnum emergens*).[15]

The first instance court rejected the claim for damages with regard to both the negative and the positive contract interest sought,[16] despite establishing that the contracting authority had failed to meet its documentary obligations.[17] The Court of Appeal identified a legal issue requiring the interpretation of the Remedies Directive and, in particular, Article 2(1)(c), which requires member states to grant their review bodies or courts the power to 'award damages to persons harmed by an infringement' of EU public procurement rules. The Court of Appeal addressed six questions on the interpretation of that provision and its interaction with the Norwegian domestic rules on damages to the EFTA Court.

The relevant questions sought clarification on how to apply the general requirement for a 'substantial breach' of EU public procurement law in the context of claims for damages, as the referring court expressed its difficulty in reconciling the CJEU judgments in *Commission v Portugal*[18] and *Strabag* with the same court's judgment in *Spijker*.[19] The questions and sub-questions thus concerned whether liability under the Remedies Directive was conditional: (i) upon the contracting authority having deviated markedly from a justifiable course of action, (ii) upon it having incurred a material error that justified a finding of culpability under a general assessment; or (iii) upon it having incurred in an inexcusable 'material, gross and obvious error' (question 1), or whether liability could be triggered under a test of 'sufficiently qualified breach' where the contracting authority was left with no discretion as to how to interpret or apply the infringed substantive rule (question 2).[20]

The EFTA Court delivered its advisory opinion in *Fosen-Linjen I* (discussed in detail below at 2.1), where it essentially found that a:

> . . . simple breach of public procurement law is in itself sufficient to trigger the liability of the contracting authority to compensate the person harmed for the damage incurred, pursuant to Article 2(1)(c) of the Remedies Directive, provided that the other conditions for the award of damages are met including, in particular, the existence of a causal link.[21]

This was contested in the context of a further appeal to the Norwegian Supreme Court, after the Court of Appeal interpreted the EFTA Court's judgment in *Fosen-Linjen I* to the effect that any breach of public procurement law is in itself sufficient for there to be a basis of liability for the 'positive contract interest', but decided not to follow it because the Court of Appeal considered that there are diverging views on this issue throughout the EEA, and that the EFTA Court's judgment in *Fosen-Linjen I* 'does not appear to be clearly correct' on that point. On that basis, the Court of Appeal decided that damages for the 'positive contract interest' are contingent on there being a sufficiently serious breach, as established by the Supreme Court of Norway in its earlier case law. The Court

14 *Fosen-Linjen I* (n 4) para 31.

15 Ibid para 32.

16 Ibid para 33.

17 Ibid para 34.

18 Judgment of 14 October 2004, *Commission v Portugal*, C-275/03, EU:C:2004:632 (not available in English).

19 *Fosen-Linjen I* (n 4) para 37.

20 Ibid para 36.

21 Ibid para 82.

of Appeal considered that this was in line not only with the CJEU judgment in *Spijker*, but also with the *Francovich* doctrine.[22]

The Supreme Court of Norway decided to request a fresh advisory opinion from the EFTA Court. It referred a single question, asking whether Article 2(1)(c) of the Remedies Directive required that any breach of the rules governing public procurement in itself is sufficient for there to be a basis of liability for positive contract interest.[23]

The EFTA Court delivered its advisory opinion in *Fosen-Linjen II* (discussed in detail below at 2.2), where it found that: 'Article 2(1)(c) of the Remedies Directive does not require that any breach of the rules governing public procurement in itself is sufficient to award damages for the loss of profit to persons harmed by an infringement of EEA public procurement rules.'[24]

At the time of writing (17 September 2019), the final decision of the Supreme Court of Norway is pending. However, the outcome of the case is not relevant for the purposes of our discussion.

2 The judgments

2.1 *Fosen-Linjen I*

In *Fosen-Linjen I*, the EFTA Court decided to group questions 1 and 2 and to carry out a joint assessment of the threshold and conditions for liability in damages for breach of EU/EEA procurement law.[25] After revisiting the goals of the then applicable EU/EEA rules (Directive 2004/18/EC),[26] the EFTA Court engaged in an analysis of the goals of the Remedies Directive.

The Court stressed that the Directive aims at 'providing adequate remedies that ensure compliance with the relevant EEA provisions on public contracts',[27] and that another of its fundamental objectives is 'to create the framework conditions under which tenderers can seek remedies in the context of public procurement procedures, in a way that is as uniform as possible for all undertakings active on the internal market. Thereby … equal conditions shall be secured.'[28]

The Court then proceeded to establish that Article 2(1)(c) of the Remedies Directive states that the measures taken concerning such review procedures must include provision for powers to award damages to those harmed by an infringement of public procurement law, but that neither such provision nor any other provisions of the Remedies Directive lay down any conditions for the award of damages as a remedy in the field of public procurement.[29] Indeed, the Remedies Directive in no way indicates that the infringement of the public procurement legislation liable to give rise to a right to damages in favour of the person harmed should have specific features.[30] In the absence of such detailed

22 *Fosen-Linjen II* (n 1) para 32.

23 Ibid para 36.

24 Ibid para 121.

25 *Fosen-Linjen I* (n 4) para 48.

26 Directive 2004/18/EC of the European Parliament and of the Council of 31 March 2004 on the coordination of procedures for the award of public works contracts, public supply contracts and public service contracts [2004] OJ L134/114. *Fosen-Linjen I* (n 4) paras 61-65.

27 *Fosen-Linjen I* (n 4) para 66.

28 Ibid para 67.

29 Ibid para 69.

30 Ibid para 71.

conditions, the EFTA Court found that it is for the legal order of each EEA state to determine the criteria on the basis of which harm caused by an infringement of EEA law on the award of public contracts must be assessed, as long as it complies with the principles of equivalence and effectiveness.[31]

The EFTA Court then restated relevant CJEU case law to the effect that the remedy of damages provided for in Article 2(1)(c) of the Remedies Directive can only constitute a procedural alternative compatible with the principle of effectiveness where the possibility of damages is no more dependent on a finding that the contracting authority is at fault than the other legal remedies provided for in Article 2(1) – such as setting aside unlawful decisions.[32] Article 2(1)(c) of the Remedies Directive therefore precludes national legislation which makes the right to damages for an infringement of public procurement law conditional on that infringement being culpable.[33] On the basis of this premise, and without much more by way of an explanation, the EFTA Court proceeded to argue that:

> . . . [t]he same must apply where there exists a general exclusion or a limitation of the remedy of damages to only specific cases. This would be the case, for example, if only breaches of a certain gravity would be considered sufficient to trigger the contracting authority's liability, whereas minor breaches would allow the contracting authority to incur no liability.[34]

The EFTA Court sought to strengthen this reasoning by arguing that 'a rule requiring a breach of a certain type or gravity would, ultimately, substantially undermine the goal of effective and rapid judicial protection sought by the Remedies Directive. It would also interfere with the objectives pursued by EU/EEA procurement law.[35] The Court added that:

> . . . the gravity of a breach of the EEA rules on public contracts is irrelevant for the award of damages. Moreover, it is not decisive for the award of damages pursuant to Article 2(1)(c) of the Remedies Directive, whether the breach of a provision of public procurement law was due to culpability and conduct deviating markedly from a justifiable course of action, or whether it occurred on basis of a material error, or whether it is attributable to the existence of a material, gross and obvious error.[36]

The EFTA Court thus answered the first two questions referred by the Court of Appeals by finding that:

> . . . the award of damages according to Article 2(1)(c) of the Remedies Directive does not depend on whether the breach of a provision of public procurement law was due to culpability and conduct deviating markedly from a justifiable course of action, or whether it occurred on basis of a material error, or whether it is attributable to the existence of a material, gross and obvious error. A simple breach of public procurement law is in itself sufficient to trigger the liability of the contracting authority to compensate the person harmed for the damage incurred, pursuant to Article 2(1)(c) of the Remedies Directive.[37]

31 Ibid para 70.

32 Ibid para 75, with reference to *Strabag* (n 7) para 39.

33 *Fosen-Linjen I* (n 4) para 77, also with reference to *Strabag* (n 7) para 45 and to *Commission v Portugal* (n 18) para 42.

34 *Fosen-Linjen I* (n 4) para 77.

35 Ibid para 79.

36 Ibid para 80.

37 Ibid para 82.

2.2 *Fosen-Linjen II*

In *Fosen-Linjen II*, the EFTA Court was asked by the Norwegian Supreme Court to essentially revisit its earlier opinion in *Fosen-Linjen I*, although formally only in relation to the liability for positive contract interest. The legal issues to be determined were essentially the same. However, their framing by the EFTA Court was rather different.

While in *Fosen-Linjen I* the EFTA Court stressed that the Remedies Directive sought 'to create the framework conditions under which tenderers can seek remedies . . . in a way that is as uniform as possible . . . [and that] equal conditions shall be secured'[38] across the internal market, in *Fosen-Linjen II*, the Court stressed that:

> . . . one of the aims of the Directive is to ensure that adequate procedures exist in all the Member States to permit the setting aside of decisions taken unlawfully and compensation of persons harmed by an infringement. Adequate review procedures . . . must not necessarily be homogenous or identical, they must merely satisfy minimum conditions, which are required by the Directive in order to ensure compliance with EEA law . . . the Remedies Directive is an instrument of minimum harmonisation.[39]

The EFTA Court then recovered the *Fosen-Linjen I* findings that 'neither Article 2(1)(c) nor any other provision of the Remedies Directive lays down specific conditions for the award of damages, which encompass specific heads of damage and the standard of liability in particular',[40] and that:

> . . . in the absence of EEA rules governing the matter, it is for the legal order of each EEA State, in accordance with the principle of the procedural autonomy of the EEA States, to determine the criteria on basis of which harm caused by an infringement of EEA law in the award of public contracts must be assessed . . .[41] [including] the criteria on the basis of which damage for loss of profit arising from an infringement of EEA law on the award of public contracts is determined and estimated, provided that the principles of equivalence and effectiveness are respected.[42]

In contrast with *Fosen-Linjen I*, however, in its second judgment, the EFTA Court sought to complement the gaps in the specific rules of the Remedies Directive with the general principle of state liability for breaches of EU/EEA law (the *Francovich* doctrine) before proceeding to assess the regulatory space left to member states. Indeed, the EFTA Court stressed that, while the standard of liability is not harmonised by the Remedies Directive:

> . . . according to the principle of State liability, an EEA State may be held responsible for breaches of its obligations under EEA law when three conditions are met: firstly, the rule of law infringed must be intended to confer rights on individuals and economic operators; secondly, the breach must be sufficiently serious; and, thirdly, there must be a direct causal link between the breach of the obligation resting on the state and the damage sustained by the injured party.[43]

The EFTA Court then proceeded to recall that compliance with the principle of effectiveness requires, in particular, that national rules cannot subject the award of

38 Ibid para 67.

39 *Fosen-Linjen II* (n 1) para 109.

40 Ibid para 111.

41 Ibid para 113.

42 Ibid para 114.

43 Ibid para 117.

damages to a finding and proof of fault or fraud.[44] However, the Court also clarified that this does not mean that certain objective and subjective factors connected with the concept of fault under a national legal system cannot be relevant in the assessment of whether a particular breach is sufficiently serious. However, the obligation to make reparation for loss or damage caused to individuals cannot depend on a condition based on any concept of fault going beyond that of a sufficiently serious breach of EEA law.[45] The EFTA Court also stressed that the requirement of a sufficiently serious breach as a minimum standard is considered sufficient for the purposes of safeguarding the rights of individuals.[46]

The EFTA Court thus answered the question referred by the Supreme Court of Norway by stating that Article 2(1)(c) of the Remedies Directive does not require that any breach of the rules governing public procurement in itself is sufficient to award damages for the loss of profit to persons harmed by an infringement of EEA public procurement rules.[47]

3 Case analysis

Although there are more differences between the two *Fosen-Linjen* judgments, it seems that the starkest contrast between them surrounds two interrelated issues. First, their approach to the level of harmonisation that the Remedies Directive imposes, where *Fosen-Linjen I* implicitly considers the Directive as requiring maximum harmonisation (i.e. equal conditions), while *Fosen-Linjen II* explicitly recognises the character of the Directive as an instrument of minimum harmonisation.[48] Second, the relationship they envisage between the general principle of state liability for breaches of EU/EEA law and the specific system of the Remedies Directive. While *Fosen-Linjen I* seems to treat the rules of the Remedies Directive in complete isolation from the general principle of state liability, *Fosen-Linjen II* explicitly establishes the link between both sets of rules.

As mentioned in the introduction, these fundamental discrepancies between the two episodes of the *Fosen-Linjen* saga are reflective of broader academic discussions underpinned by a perceived inconsistency in the CJEU case law in this area and, in particular, the judgments in *Strabag* and *Spijker*, which are discussed in minute detail in both *Fosen-Linjen* judgments. Moreover, the exact same arguments had already been considered by the UK Supreme Court in 2017. This section looks in more detail at both issues.

3.1 Procurement remedies and state liability: unitary vs separation theses

One of the distinctive features of EU public procurement law as compared to more general EU internal market law lies in its specific enforceability through the mechanisms regulated in the Remedies Directive, of which the 'aim is to make more concrete the obligations imposed by EU law, in the hope that this will improve policing of the rules governing contracts awarded by public authorities and in consequence develop further

44 Ibid para , with reference to *Fosen-Linjen I* (n 4) para 75, as well as to *Strabag* (n 7) paras 39–40 and *Commission v Portugal* (n 18) para 42.

45 *Fosen-Linjen II* (n 1) para 119.

46 Ibid para 120.

47 Ibid para 121.

48 For extended discussion from the perspective of minimum harmonisation, see A Sanchez-Graells, 'The EFTA Court's *Fosen-Linjen* Saga on the Liability Threshold for Damages Claims for Breach of EU Public Procurement Law: A There and Back Again Walk' (2019) European Procurement and Public Private Partnership Law Review (forthcoming) <https://ssrn.com/abstract=3455213>.

the construction of a true internal market in this sector'.[49] As we have seen, however, the relationship between the Remedies Directive and the general rules on national remedies for breaches of EU law is contested.

The existing debate about the relationship between these two regulatory mechanisms boils down to disagreements over whether the Remedies Directive should be constructed as a particularisation of the general principle of state liability under EU law (a 'unifying thesis') or whether a distinction should be made between 'a public law of torts in the form of Member State liability, and damages for breaches of specific EU legislation under the effectiveness postulate (the "separation thesis")'.[50] The unifying thesis would result in the superimposition of the requirement of 'sufficiently serious breach' to the award of damages under the Remedies Directive. Conversely, the separation thesis would result in a free-standing interpretation of the liability threshold in the Remedies Directive and, possibly, in a reduction of the threshold of infringement triggering potential liability in damages for 'any infringement'. This would aim to avoid what has been considered 'the paradoxical result . . . that although the remedies regime is more concrete and elaborate than in other areas of the law, the Court [of Justice] would be forced into the abstract generalities of Member State liability, rather than the specificities of the procurement sector'.[51] Moreover, the separation thesis sometimes receives support on the basis that the Remedies Directive pre-dates the seminal case of *Francovich*, which opens up arguments about a possible divergence in the legal tests and, specifically, in the liability threshold applicable to each of these heads of claims for damages against the state.

3.2 Perceived inconsistencies in CJEU case law

Most of the legal arguments submitted to the EFTA Court in the *Fosen-Linjen* saga related to the consistency or incompatibility of the CJEU's judgments in *Strabag* and *Spijker*. I argue that any perceived inconsistencies result from an incorrect conceptualisation of the regulation of the system of procurement remedies and its interpretation by the CJEU.

A preliminary point needs to be made to stress that the interaction between the Remedies Directive and the general doctrine of state liability for breaches of EU law has been clarified by the CJEU. Indeed, this was explicitly addressed in *Spijker*,[52] when the CJEU stated that Article 2(1)(c) of the Remedies Directive '*gives concrete expression to the principle of State liability* for loss and damage caused to individuals as a result of breaches of EU law for which the State can be held responsible'.[53] Moreover, the CJEU was clear that:

> … as regards state liability for damage caused to individuals by infringements of EU law for which the state may be held responsible, *the individuals harmed have a right to redress where* the rule of EU law which has been infringed is intended to confer rights on them, *the breach of that rule is sufficiently serious*, and there is a direct causal link between the breach and the loss or damage sustained by the individuals. In the absence of any provision of EU law in that area, *it is for the internal legal order of each member state, once those conditions have been complied with, to determine the criteria on the basis of which the damage arising from an infringement of EU*

49 S Weatherill, 'EU Law on Public Procurement: Internal Market Law Made Better' in X Groussot, J Hettne and S Bogojevic (eds), *Law and Discretion in EU Public Procurement*, vol 26, Studies of the Oxford Institute of European and Comparative Law (Hart 2019) 21, at 43.

50 H Schebesta, *Damages in EU Public Procurement Law* (Springer 2016) 8. For extended discussion, see ibid 65–71, esp 67–68.

51 Ibid 71.

52 *Spijker* (n 8).

53 Ibid para 87, emphasis added.

> *law on the award of public contracts must be determined* and estimated, provided the principles of equivalence and effectiveness are complied with.[54]

There could not be a closer formulation of the unifying thesis than in *Spijker*,[55] whereby it is clear that Article 2(1)(c) of the Remedies Directive fleshes out or particularises the doctrine of state liability for breaches of EU law in the context of public procurement. As mentioned above, some objections could be raised to the effect that, the Remedies Directive having been adopted in 1989, it could not have logically given expression to the principle of state liability for breach of EU law, as it was only formulated in *Francovich* in 1991. However, such objections can be dismissed on the basis of different types of arguments. A practical argument is that the Remedies Directive was revised in 2007, when the principle of state liability was already consolidated in CJEU case law, and the EU legislator did not consider it necessary to make any changes to Article 2(1)(c). A jurisprudential argument could also be used to dismiss the objection, on the basis that the CJEU does not create general principles of EU law in its case law, but rather draws from them or declares them – which logically requires their pre-existence (arguably, from the origins of the Treaties).[56] Ultimately, at least from a positive perspective, the existence of the link between the general principle and the Directive could (and should) be allowed to rest on the simple fact that the CJEU has exclusive competence to carry out such interpretation *ex* Article 267 Treaty on the Functioning of the EU.[57]

Spijker is, however, not universally seen as having settled the issue of the interaction between the grounds for actions for damages under the Remedies Directive and under the state liability doctrine. The grounds for rejection of the unitary thesis are usually found in additional CJEU case law – namely, in the *Strabag* Judgment – which barred the possibility of subjecting the liability in damages of a contracting authority to a requirement of fault or fraud,[58] even if claimants can benefit from a rebuttable presumption of fault.[59] Some authors consider the unitary thesis irreconcilable with a reading of *Strabag* that would require member states to ensure *strict liability* for breaches of EU public procurement law. However, I would argue that this conflates the aspect of objectivity of the assessment with the distinct issue of the threshold for liability (as not every objective breach is necessarily sufficiently serious and thus not every breach needs to result in liability).[60]

Those readings of *Strabag* are incorrect in that they miss the different levels of regulatory design at which *Spijker* (top layer) and *Strabag* (second layer) operate.[61] In other words, *Spijker* establishes a link between the general principle of state liability and the Remedies Directive, which gives it concrete expression. Within this framework, *Strabag* (and other case law) modulates the requirements of the principles of effectiveness and equivalence that result from the higher regulatory layer. In any case, the main point of

54 Ibid para 92, emphasis added.

55 In agreement on the positive description, but criticising it normatively, see Schebesta (n 50) 65–72.

56 This is an issue that, however, exceeds the scope of this article and, consequently, will not be assessed in any detail.

57 Along the same lines, H C H Hofmann, 'The Court of Justice of the European Union and the European Administrative Space' in M W Bauer and J Trondal (eds), *The Palgrave Handbook of the European Administrative System* (Palgrave 2015) 301–12.

58 See also *Commission v Portugal* (n 18).

59 *Strabag* (n 7).

60 For discussion, see A Sanchez-Graells, 'Assessing the Public Administration's Intention in EU Economic Law: Chasing Ghosts or Dressing Windows?', in K A Armstrong (ed), *Cambridge Yearbook of European Legal Studies 2016* vol 18 (Cambridge University Press 2016) 93–121.

61 To the same effect, see *Nuclear Decommissioning Authority* (n 9) per Lord Mance, at [24].

contention rests on what could be seen as a *lex specialis* understanding of the interaction between the two regulatory frameworks – i.e. a view that the general condition for there to be a 'sufficiently serious breach' of EU law under state liability is contrary to the requirement for strict liability for breaches of EU procurement law, which would have led the Remedies Directive to impose a lower triggering threshold by solely mentioning the need for an unqualified infringement as sufficient ground for damages claims (Article 2(1)(c)).[62] As detailed above (section 2), the latter view was reignited by the EFTA Court in *Fosen-Linjen I*, but subsequently abandoned in *Fosen-Linjen II*. I argue that *Fosen-Linjen II* represents the right statement of EU law in this area, in particular in view of the minimum harmonisation carried out by the Remedies Directive, as discussed in the following section. Remarkably, this was the position that the UK Supreme Court had already reached before the *Fosen-Linjen* saga.

3.3 The UK Supreme Court's view

In its *Nuclear Decommissioning Authority* judgment,[63] the UK Supreme Court followed what I think is the correct reading of *Spijker* and established that it makes clear:

> . . . that the *liability of an awarding authority is to be assessed by reference to the Francovich conditions*. Subject to these conditions being met . . . [it goes] on to make clear that the criteria for damages are to be determined and estimated by national law, with the further caveat that the general principles of equivalence and effectiveness must also be met . . . Finally, [it] summarises what has gone before, repeating the need to satisfy the Francovich conditions.[64]

More importantly, the UK Supreme Court considered that:

> . . . there is . . . very clear authority of the Court of Justice confirming that the liability of a contracting authority under the Remedies Directive for the breach of the [public procurement rules] . . . is in particular only required to exist where the minimum Francovich conditions are met, although it is open to States in their domestic law to introduce wider liability free of those conditions.[65]

Therefore, along the same lines of *Fosen-Linjen II*, the UK Supreme Court followed a unifying thesis compatible with minimum harmonisation and took the clear view that *as a matter of EU law* the existence of grounds for an action in damages based on the Remedies Directive requires the existence of a 'sufficiently serious breach' of EU public procurement law. The UK Supreme Court explicitly ruled out any inconsistency between this approach and other case law of the CJEU, in particular *Strabag*, on the basis that the cases are not incompatible and, importantly, that the CJEU 'in Spijker was aware of the recent decision in [*Strabag*], cited it . . . and clearly did not consider it in any way inconsistent with what [it] said about the general applicability of the Francovich conditions'.[66] Importantly, the UK Supreme Court took no issue with the possibility for more generous domestic grounds for actions for damages.[67] On the whole, the UK Supreme Court considered that 'there is no uncertainty or confusion in the Court of

62 Whether this is compatible with a unifying thesis or with a separation thesis, or neither of them, remains unclear, but this aspect of the discussion exceeds the scope of this article.

63 *Nuclear Decommissioning Authority* (n 9) as per Lord Mance, with Lady Hale and Lords Neuberger, Sumption and Carnwath agreeing.

64 Ibid per Lord Mance, at [23], emphasis added.

65 Ibid at [25], emphasis added.

66 Ibid at [24], emphasis added.

67 Although it eventually decided that this was not the case in relation to the Public Contract Regulations 2006; see *Nuclear Decommissioning Authority* (n 9) per Lord Mance, at [37], with which I also agree.

Justice's case law, and that [it is safe to rely] on the clear language and ruling in *Spijker* as settling the position, whatever may have been previous doubts or differences of view at national level'.[68]

Conclusion

The controversy underpinning the *Fosen-Linjen* saga is ultimately reflective of discussions about the interrelation between the Remedies Directive and the more general principle of state liability for breaches of EU/EEA. The CJEU had unequivocally established in 2010 that the Remedies Directive '*gives concrete expression to the principle of State liability* for loss and damage caused to individuals as a result of breaches of EU law for which the State can be held responsible'.[69] Therefore, only 'sufficiently serious breaches' of EU/EEA procurement law give rise to liability in damages, provided the other requirements are present. This should have put the discussion to rest, as forcefully argued by the UK Supreme Court in 2017.[70] However, as *Fosen-Linjen I* evidenced, the debate lingered on. The reversal of the EFTA Court's view in *Fosen-Linjen II* should serve to bury the issue. However, this may encounter some additional resistance. Therefore, the CJEU would be well advised to (re)confirm the unitary thesis enshrined in *Spijker* at the earliest opportunity to avoid any perpetuation of this debate in the context of EU/EEA public procurement law.

Postscript

The Norwegian Supreme Court delivered its judgment in the so-called *Fosen-Linjen* case on 27 September 2019 (HR-2019-1801-A). As a general approach, the Norwegian Supreme Court followed the unitary thesis reflected in *Fosen-Linjen II*. However, it granted the appellant compensation for the negative interest, but not for the positive interest. This raises a number of interpretive issues concerning the boundary between the liability threshold and the analysis of causation in claims for damages due to a breach of procurement rules. The interested reader will find detailed analysis in D S Lund, 'The Fosen-Linjen Saga – a Norwegian Perspective' (2019) European Procurement and Public Private Partnership Law Review (forthcoming).

68 Ibid at [26], with reference to A Collins, 'Damages in Public Procurement – An Illusory Remedy?', in K Bradley, N Travers and A Whelan (eds), *Of Courts and Constitutions. Liber Amicorum in Honour of Nial Fennelly* (Hart 2014) 339.

69 *Spijker* (n 8) para 87, emphasis added.

70 *Nuclear Decommissioning Authority* (n 9).