Modern Defamation Law: Balancing Reputation and Free Expression Edited by David Capper

Queen's University Belfast

Modern Defamation Law

Balancing Reputation and Free Expression

Edited by
David Capper

Modern Defamation Law:

Balancing Reputation and Free Expression

An intense media campaign in recent years has resulted in the introduction of a new defamation Bill to the British parliament, effecting important changes to the law with a view to the protection of free expression and the curtailment of libel tourism. This book critiques these changes in the equally important context of the protection of reputation. It will be essential reading for all those whose work concerns the publication of material with any potential to raise issues of public controversy, whether in media, politics, science or academia.

ISBN: 978-1-899177-24-0

Price £10 plus £2.50 p&p

Professor Sally Wheeler, School of Law, Queen's University Belfast
BELFAST BT7 1NN

Order Form

Please send me ___ copies of *Modern Defamation Law.*

@ £10 + £2.50 p&p per copy. **Total** £ _____

Name _____

Address _____

_____

_____ Postcode _____

Please make cheques payable to Queen's University Belfast.
Send your order to:

Professor Sally Wheeler
School of Law
Queen's University Belfast
BELFAST BT7 1NN

Northern Ireland Legal Quarterly

Volume 64 Number 4

EDITOR
PROFESSOR SALLY WHEELER

Queen's University Belfast
School of Law

Contents

NILQ 64(4): 425–42

Licences of business premises: contract, context and the reach of *Street v Mountford*

MICHAEL HALEY*

Keele University

Part II of the Landlord and Tenant Act 1954 confers on those business tenants within its remit the primary right to apply to court for the grant of a new lease at a market rent.[1] The laudable ambition is, as Lord Wilberforce explained, 'to provide security of tenure for those tenants who had established themselves in business in leasehold premises so that they could continue to carry on their business there'.[2] In marked contrast with the Rent Act 1977 and the Housing Act 1980, the Part II provisions do not constitute a tenants' charter.[3] The restrictions imposed by Parliament are, therefore, of comparatively modest design and when, as in recent times, market forces favour the tenant, these become of much reduced importance. Seemingly distant are the days when, as Briggs LJ acknowledged, 'spiralling property prices meant that unscrupulous private landlords could reap large profits'.[4] Due to this coincidence of factors, the business tenancy code has proved resilient to shifts in political and social ideology.[5] Unlike its agricultural and residential counterparts, it has managed to evade the culling effect of deregulation and now 'stands alone as a code of major practical significance involving restrictions on the landlord's freedom of contract'.[6] The prevailing sentiment is that the 1954 Act has functioned satisfactorily[7] and the official rhetoric remains supportive of the principle that, 'business tenants should normally have a right to renew their tenancies'.[8] The underlying irony, however, is that since 1969 Parliament has innovated reforms that fundamentally dilute the impact of regulatory

* Professor of Property Law, Keele University.

1 For tenants who are unable to obtain a renewal, flat-rate compensation might be available under s 37 to help them re-establish their business elsewhere.

2 *O'May v City of London Real Property Co Ltd* [1983] 2 AC 726, 747.

3 See M Haley, 'The Statutory Regulation of Business Tenancies: Private Property, Public Interest and Political Compromise' (1999) 19 Legal Studies 207.

4 *Lambeth LBC v Loveridge* [2013] EWCA Civ 494, [1].

5 See R J Heffron and P Haynes, 'Striking a Statutory Balance: Constant Change in Residential Leases versus Static Change in Commercial Leases' (2012) Conv 195.

6 Law Reform Advisory Committee for Northern Ireland, *A Review of the Law relating to Business Tenancies in Northern Ireland* (Discussion Paper No 3, 1992) [2.2.5].

7 See, for example, Law Commission, *A Periodic Review of the Landlord and Tenant Act 1954 Part II* (Law Com No 208, 1992); Law Commission, *Part II of the Landlord and Tenant Act 1954* (Working Paper No 111, 1988) and the conclusions of the Department of the Environment, HC Deb 20 November 1985, vol 87, col 245.

8 Department for Communities and Local Government, *Landlord and Tenant Act 1954: Review of Impact of Procedural Reforms* (DCLG Report 2006) [4].

control. Such developments as the introduction and further facilitation of contracting-out[9] and the ability to grant short leases that fall beyond the reach of statutory protection[10] ensure that the Part II machinery has become optional at the behest of the parties.

Within this relaxed legal framework, it might be assumed that the traditional utilisation of licence agreements as an avoidance measure would fall into desuetude.[11] This is particularly so in the aftermath of *Street v Mountford*,[12] which effectively sounded the death knell for the private sector residential licence.[13] Some 30 years on, there has been a steep decline in judicial and academic commentary concerning the lease or licence distinction. This period of quietude might suggest that licences are similarly moribund in the commercial sector. Licences of commercial premises are, however, enjoying a renaissance, being the preferred choice of many business occupiers. This unusual state of affairs is due to prevailing market forces, which have for some years disadvantaged the landowner and yet sowed doubt as to the long-term future of many businesses. The licence offers the perfect compromise. For the landowner, the risk of holding a portfolio of vacant commercial properties is reduced and, from the occupier's perspective, the dangers inherent in being tied to a long-term relationship are avoided. Despite this market revitalisation, however, the law remains in an unpredictable and thoroughly unsatisfactory state, out of kilter with modern reality. It is to be expected that any recasting of a commercial licence agreement would be premised upon neutral and coherent legal reasoning, fashioned to the policy and mischief underpinning the legislation to be sidestepped. Similarly, it is to be hoped that there is some cogent and compelling justification for an outcome that allows a commercial entity or, indeed, third party[14] to unravel a bargain voluntarily entered. Nevertheless, the prevailing judicial approach to the construction of commercial sector licence agreements demonstrates no such overarching traits.

The present state of the law is directly attributable to Lord Templeman's speech in *Street v Mountford* and the manner in which it has been applied in the lower courts.[15] As will become clear, this decision unfolds on three levels. First, it acknowledges that exclusive possession is the vital indicator of a tenancy. Secondly, it emphasises that the substance of the agreement must triumph over its outward form. Thirdly, it encourages the judiciary to be 'astute to detect and frustrate sham devices and artificial transactions whose only object is to disguise the grant of a tenancy'.[16] The latter advice was a deliberate attempt to staunch the flow of licences of residential property. Uncertainty has arisen, however, as to the applicability of this tripartite approach within the commercial context. Some judges have applied it slavishly, failing to prioritise explicitly the context in which the agreement arises.[17] Such cases often involve a shift of emphasis from construction to the reconstruction of contractual bargains.

9 S 38A of the 1954 Act allows contracting-out by private arrangement following the giving of an informational 'health warning' to the tenant and the tenant's timely declaration of agreement.

10 S 43(3).

11 Licence agreements fall totally outside the Part II framework: *Shell-Mex & BP Ltd v Manchester Garages* [1971] 1 WLR 612.

12 [1985] 1 AC 809.

13 See P Sparkes, 'Breaking Flat Sharing Licences' (1989) 5 JSWL 293. The advent of the landlord-friendly assured shorthold tenancy in the Housing Act 1988 also made residential licence agreements unnecessary.

14 For example, HMRC in *Grimsby College Enterprises Ltd v Revenue & Customs Commissioners* [2010] UKUT 141 and the freeholder in *London Development Agency v Mehmet* [2009] EWHC 1730 (Ch).

15 See P F Smith, 'A Dent in *Street v Mountford*' (1987) Conv 220.

16 *Street v Mountford* (n 12) 825.

17 A myopia demonstrated by Deputy Judge Paul Baker QC in *London & Associated Investment Trust plc v Calow* (1987) 53 P&CR 340, 352: 'I cannot see any differentiation for this purpose between a residence on the one hand and a solicitor's office on the other.'

Others, most notably Arden LJ in *NCP Ltd v Trinity Development Co (Banbury) Ltd*,[18] have paid it little more than lip service, preferring instead to adopt a results-led style of reasoning. This article will argue that the latter of these diametrically opposed standpoints is correct and that there should be no roving commission to determine the existence or non-existence of a licence of commercial premises.

Although the traditional perception is that licence agreements are employed to the disadvantage of the occupier, this article will challenge that preconception. Unlike in the housing sector, for commercial licensees there are real attractions and substantial benefits associated with a licence of business premises.[19] From this perspective, it will be argued that the *ratio* of *Street* is properly to be confined to the residential sector. It is there that the mischief it addressed is localised and where there is acute need 'to protect a person whose bargaining power is handicapped by his personal circumstances'.[20] While the protective impulse to seek out artifice in the housing context has undeniable socio-economic merit,[21] the investigative stance and forensic approach promoted by Lord Templeman are unsuitable for universal application.[22] Within the commercial sector, there is no scope for the interventionist ideal[23] and paternal sentiments are misplaced, unnecessary and counterproductive. As Bridge commented: 'one would have thought that with the greater awareness of the full implications of the bargain which comes with legal advice and business knowhow, the courts can afford to be less paternalistic and leave them to sort out their rights and duties between themselves'.[24]

A persistent failure of the lower courts to acknowledge this difference, however, entails that the crucial intersection between contract and regulatory context is overlooked. Despite judicial protestations to the contrary,[25] even a cursory examination of the authorities reveals an ingrained predilection towards identifying a relationship of landlord and tenant.[26] As Bridge acknowledged: 'The message is clear that courts will look at such contracts with a heavy bias in favour of a tenancy.'[27]

Perceptions and perspective

The statutory regulation of the landlord and tenant relationship marked a shift from contract to status. Inevitably, this introduced new distinctions, accentuated existing differences and emphasised definition and function as dictating the degree of statutory protection afforded (if any). It is in this context that the demarcation between a lease of land and a licence to occupy assumes significance. Comprehensive statutory protection has throughout been restricted to qualifying tenants. As Bridge explained, 'the existence of a

18 [2002] 2 P&CR 18.

19 Especially so the so-called 'easy-in, easy-out' licence (as discussed below).

20 Purchas LJ in *Crancour Ltd v Da Silvaesa* (1986) 52 P&CR 204, 227.

21 See P Vincent-Jones, 'Exclusive Possession and Exclusive Control of Private Rented Housing: A Socio-Legal Critique of the Lease–Licence Distinction' (1987) 14 Journal of Law and Society 445.

22 There is, as Deputy Judge Richard Southwell QC acknowledged in *Mehta v Royal Bank of Scotland* (2000) 32 HLR 45, 53, 'no simple all embracing test . . . The search for such a test would be a search for a chimaera'.

23 In *Clear Channel UK Ltd v Manchester CC* [2006] 1 EGLR 27, [28], Jonathan Parker LJ lamented the licensee's attempt to overturn its licence agreement, deriding it as, 'surprising and . . . unedifying'.

24 S Bridge, '*Street v Mountford* – No Hiding Place' (1986) Conv 344, 351. This view is evidently shared by Arden LJ in *NCP v Trinity Development* (n 18) [29], who, by way of a brave departure from traditional constraints, was determined to uphold a commercial licence, seemingly at the expense of legal principle.

25 See, for example, Buckley LJ in *Shell-Mex & BP* (n 11) 619.

26 See *Pankhania v London Borough of Hackney* [2002] EWHC 2441 (Ch), where (at [11]) Deputy Judge Rex Tedd QC found it to be 'an inexorable inference' that exclusive possession had been granted in relation to a car park.

27 Bridge (n 24) 352.

tenancy is the *sine qua non* of the protected breed'.[28] Licence agreements are dismissed as too personal, transient and precarious[29] to be the subject of full legal protection. Hence, such informal arrangements represent the traditional and most convenient method by which the statutory rights of the occupier can be bypassed.[30]

On the scale of legislative intervention, the protection of the family home consistently ranks highest in priority. It is here that the inequality of bargaining power between the parties is most pronounced and the countervailing social rights of property in more urgent demand.[31] Unsurprisingly, the greater the degree of protection afforded the more enterprising and imaginative are the attempts to take advantage of loopholes existing within the relevant legislative code.[32] Hence, 'the mad maelstrom of the Rent Acts'[33] afforded the battleground on which the majority of the lease or licence cases, including *Street v Mountford*, have long been fought. It also offers the prime vantage point from which the case-law concerning the identification of avoidance and evasion may properly be understood and the sins of landlordism fully exposed.[34]

As mentioned, and unlike in the residential sector, the parties can directly contract out of Part II of the Landlord and Tenant Act 1954. Hence, landlords of commercial property can already grant long fixed-term tenancies that afford no renewal or compensation entitlements.[35] There is no need to utilise a licence to achieve the same outcome. Indeed, and again very unlike the housing rental market, the conventional wisdom is that such long leases better serve the interests of the commercial landowner. The licence does not, therefore, have innate appeal for many commercial landlords and, most certainly, is not appraised simply as an avoidance measure. It is a device that, in the current economic climate, predominantly serves the interests of the occupier. Nevertheless, for some property owners there may still be discernible advantages. The landowner's plans may be uncertain, for example, it might be debating whether to sell, redevelop or occupy the property itself for business purposes. A licence would amount to an eminently sensible arrangement in circumstances where the occupier moves in under the auspices of an agreement for lease, pending the grant of the lease itself[36] or the occupier is to have a trial period as licensee with an option to take a lease.[37] The landowner might, moreover, be awaiting a change in market forces before committing to a more permanent relationship. In the context of the statutory renewal machinery, the creation of a licence could be necessary to ensure that the

28 S Bridge, 'Down to Business with *Street v Mountford*' (1987) Conv 137, 137.

29 Licence agreements cannot bind third parties: *London Development Agency v Mehmet* (n 14).

30 This is not to say that licensees are denied all statutory rights. For example, the Agricultural Holdings Act 1986 treats certain licences as if they are tenancies; see also ss 1–3 of the Protection from Eviction Act 1977; ss 76–85 of the Rent Act 1977; and s 79 of the Housing Act 1985.

31 As Lord Templeman observed in the conjoined appeals *AG Securities v Vaughan; Antoniades v Villiers* [1990] 1 AC 417, 458, 'in a state of housing shortage a person seeking residential accommodation may agree to anything to obtain shelter'.

32 For example, as a means of outflanking the Rent Act 1977 landlords foisted inner city 'holiday' lets and unwanted board and attendances upon occupiers.

33 Danckwerts LJ in *Myer v Mercantile Properties Ltd* (1961) 179 EG 693, 695. R Street, 'Coach and Horses Trip Cancelled? Rent Act Avoidance after *Street v Mountford*' (1985) Conv 328, 328, complained (after his defeat in the House of Lords) that: 'The Rent Acts are grossly unfair to landlords.'

34 As the Law Commission report, *Landlord and Tenant: Reform of the Law* (Law Com No 162, 1987) [4.14] accepted: 'It is all the more deplorable when it relates to rules which are intended to protect the vital interests of individuals who include some of the most vulnerable in our society.'

35 This possibility extends to all fixed-term tenancies: see M Haley, 'Contracting Out and the Landlord and Tenant Act 1954: The Ascendancy of Market Forces' (2008) Conv 281.

36 *Cameron Ltd v Rolls-Royce plc* [2007] EWHC 546 (Ch).

37 *Essex Plan Ltd v Broadminster* (1988) 56 P&CR 353.

licensor (in the guise of mesne landlord) itself retains renewal rights.[38] If the licensor holds as tenant under a lease that prohibits assignment and sub-letting, the licence agreement will represent the only way in which the tenant can lawfully part with possession. Similarly, a tenancy proper cannot exist where the licensor itself occupies under a licence agreement.[39] The ability to deny a licensee exclusive possession might also bestow practical advantages by allowing the licensor to reserve more extensive rights over the premises. This might be crucial where there is to be occupation of, say, a stall in an indoor market or a retail concession in a theatre or department store.[40] In the process, the licensor might be able to prevent the arrangement being treated as 'leasing or letting of immovable property' for VAT purposes.[41] A denial of exclusive possession will enable the licensor to move the licensee to other parts of the premises, require the licensee to share the accommodation with others and control such mundane matters as signage, advertising and opening hours.[42] The licensor can, therefore, legitimately retain a degree of territorial control that is anathema to the existence of a tenancy and totally unrealistic in the residential context.[43]

Although the licence is conventionally appraised as a means of promoting the licensor's interests at the expense of the licensee, this is clearly not so in the commercial sector. As indicated, in this market, the licence is primarily tailored to the interests and demands of the occupier. From the licensee's viewpoint, the arrangement might be highly responsive to existing and future business plans. The benefits focus on it being a much more fluid and adaptable arrangement than the formal and heavily stylised relationship of landlord and tenant. Indeed, there is evidence that leases in the UK tend to be longer than elsewhere in the world and that many business tenants do not appreciate many of the legal implications of signing tenancy agreements.[44] It has, therefore, been argued that those taking leases for the first time 'are vulnerable to abuse given the complexity and length of many lease documents'.[45] By way of contrast, the licence will be drafted in less arcane language than that employed in a lease and will spell out more clearly the rights and obligations of the parties. The licence can, in particular, prove a highly attractive and cost-effective option for small and start-up businesses that face an uncertain future[46] and are reluctant to be tied to a more regulated and long-term relationship.[47] Such is evidenced by the market growth of the so-called 'easy-in, easy-out' licences of office space within the many business centres

38 See *Graysim Holdings Ltd v P&O Property Holdings Ltd* [1995] 4 All ER 831.

39 *London Development Agency v Mehmet* (n 14). There a company had no title to grant leases in respect of two shops so the freeholder was entitled to possession.

40 For example, in *Smith v Northside Development Ltd* (1988) 55 P&CR 164, there was a licence to share floor space in a unit at Camden Lock Market.

41 VAT is payable as regards licences where the licensee has the right to occupy that area as owner and to exclude others from enjoying that right (HMRC Reference: Notice 742, June 2012); see the failed attempt in *Grimsby College v Revenue & Customs Commissioners* (n 14).

42 See, for example, the garage forecourt cases: *Shell-Mex & BP* (n 11); *Esso Petroleum Co Ltd v Fumegrange Ltd* (1994) 68 P&CR D15, where the rights and control retained there by the licensors were, when viewed cumulatively, utterly inconsistent with the occupiers having exclusive possession of the premises.

43 *NCP v Trinity Development* (n 18) illustrates that clauses denying exclusive possession of a shoppers' car-park can be upheld whereas equivalent terms would be struck out in the residential context.

44 See the government-funded research undertaken by N Crosby, C Hughes and S Murdoch, *Do Small Business Need Protecting in Commercial Lease Negotiations?* (ERES 2005). Download available from <www.centaur.reading.ac.uk>.

45 Ibid 16. The authors add, at 1: 'Many small business tenants are unrepresented at the commercial stage of negotiations and take the first terms on offer'. They note also that these tenants also remain unaware of the voluntary industry Codes of Practice that are published from time to time.

46 For example, as regards a business that operates on fixed-term supply contracts.

47 A licence can be for an uncertain duration whereas a tenancy cannot: *Mexfield Housing Co-operative Ltd v Berrisford* [2012] 1 AC 955.

operating in the UK. These facilities range from the small owner-operated centres (frequently offered by local authorities) to large premises managed by mainstream property groups. Located in the city centre or out-of-town commercial site, they are targeted primarily at new businesses. These agreements allow the licensee to move into a serviced business centre, office block or executive suite on the payment of a licence fee (usually payable monthly) and the provision of a modest and refundable cash deposit. Legal and surveying costs incurred in relation to a licence agreement are much reduced; no capital premium will be payable and no stamp duty land tax or other registration fee will become due. This type of licence also enables the occupier to manage effectively its monthly outgoings without having to make a substantial payment to move in, to pay several months of rent in advance or to be locked into to a rent review clause.[48] Many licensors will provide office services, meeting rooms and office equipment on a pay-for-use basis. The agreement might also feature such items as security and concierge services, broadband facilities, parking rights and furnishings as part of an inclusive package. This may well generate further savings on staffing costs and initial capital outlay. Serviced office space has, therefore, become a cost-effective option that can achieve substantial long-term savings for the licensee. Such was demonstrated in a cost comparison survey undertaken by the Chartered Institute of Purchasing and Supply, which revealed that serviced space can offer savings for the occupier of up to 91 per cent.[49] The survey, as conducted in eight major cities, showed average savings of up to 50 per cent for those firms taking a licence of space in a business centre.

Although leasehold premises are readily to be found, they are much more difficult to leave. This is problematic as the evidence assembled by the Crosby, Hughes and Murdoch research shows 'that many small businesses will need to change their business premises within fairly short time frames'.[50] This might be because they outgrow the existing premises, there is a change in the market, or the area becomes unsuitable. Under a licence agreement, however, the relationship will terminate either on the date specified or agreed between the parties, or, absent agreement, by one party giving the other reasonable notice.[51] The licensee will be released from continuing liability as soon as the contract ends. There is no risk of an unwanted continuation tenancy arising under the 1954 Act nor is there any threat of forfeiture for breach of covenant. The inherent flexibility of this style of arrangement entails also that, if and when the business changes or expands, there will be the possibility of licensing further space within the building and again without the formality and costs associated with the grant of a lease. Until that time, the licensee only ever pays for the space that it needs and can upsize or downsize, while maintaining the same address, with the minimum of disruption. Not surprisingly, many businesses view this type of arrangement as ideal in the prevailing economic climate. It offers a modern, attractive and smarter means of satisfying their accommodation needs.

The need for certainty in matters commercial is of paramount importance and demands that an agreement within the business community be construed so as 'to give it the meaning and effect which both parties must have intended given the terms and structure of their contract'.[52] The factual matrix in which the agreement is sited cannot be ignored and it

48 It also shields the occupier from the self-help remedy of distress for rent as there is neither a rent payable nor a tenancy in existence.

49 *True Cost of the Flexible Office Survey* (Chartered Institute of Purchasing and Supply 2001).

50 Crosby et al (n 44) 5. The research (at 11) shows, moreover, that the practice of including break clauses in leases differs according to the size of the tenant's business.

51 In *Smith v Northside Development* (n 40), one month's notice to terminate was held to be reasonable.

52 Patten LJ in *Gavin v Community Housing Association Ltd* [2013] EWCA Civ 580, [42].

would plainly be counterintuitive for the courts to unravel this style of informal and occupier-orientated relationship. Nevertheless, under the current law there is no guarantee that a professionally styled licence agreement, freely entered into at arm's length between commercial entities, will be upheld.[53] The courts instead profess a keen ability to decipher the existence of a tenancy against the expressed intentions of the parties. Of vital importance, therefore, is what Partington described as the 'murky distinction between the "genuine transaction" and the "mere sham"'.[54] As will become apparent, the legal techniques practised by the judiciary when distinguishing between the real agreement and an inauthentic version thereof are roughly hewn and, although high in rhetoric, tend to be low on substance and principle. They also fail to discriminate between the protective measures required in the residential market and the very different demands of the commercial sector.

Avoidance or evasion?

While lawful avoidance is perfectly permissible,[55] illegitimate evasion most certainly is not.[56] Unfortunately, the boundary between the two is far simpler to state than it is to identify on any given facts. As Lord Greene MR pointed out, it is because of the 'highly technical and highly variegated relationship of landlord and tenant . . . almost inevitable that fine distinctions will be found to prevail'.[57] The standard explanation is as offered by Gavan Duffy J: 'You do not evade an Act by doing something which is not forbidden by the Act, but you do evade an Act by doing something which is prohibited under the guise of doing something else.'[58] Accordingly, evasion involves an element of deception and artifice and arises where the contract does not 'reflect the true intention of the parties in entering into an agreement in order to conceal the reality of the transaction'.[59] The court, moreover, must be 'especially careful to see that the wool is not being pulled over its eyes'[60] and, if an artificial transaction is detected, effect must be given to the concealed bargain.[61] While the language may be assuring, less convincing is the manner in which the judiciary undertakes the delicate task of distilling the true agreement from the apparent bargain. As Bridge put it, 'beyond a scintilla of a doubt is that the lease–licence issue is one of the most complex and daunting a court may have to contend with'.[62] The primary obstacle is that a tenancy agreement and a contractual licence will contain a similar package of core terms, relating to, for example, duration, termination, payment, maintenance and use. As Buckley LJ observed of the licence agreement in *Shell-Mex & BP Ltd v Manchester Garages*, 'many of the clauses in it are clauses which could appropriately find their place in a tenancy agreement . . . but it

53 See *Grimsby College v Revenue & Customs Commissioners* (n 14) (occupation of Engineering Centre and right to share equipment); *Dellneed Ltd v Chin* [1987] 1 EGLR 75 (management agreement of a restaurant); *University of Reading v Johnson-Houghton* [1985] 2 EGLR 113 (use of gallops for horse training).

54 M Partington, 'Non-Exclusive Occupation Agreements' (1979) 42 MLR 331, 331

55 It is, as Lloyd J emphasised in *Brumwell v Powys County Council* [2011] EWCA Civ 1613, [28], 'perfectly legitimate for the parties to enter into agreements the substance of which was designed to prevent that legislation from applying'.

56 It is not possible for the landlord, as Lawton LJ commented in *O'Malley v Seymour* (1983) 7 HLR 70, 81: '[To] arrange his affairs one way, which brings his property within the Rent Acts, and then dress them up in another way so as to give the impression that it is outside the Rent Acts'.

57 *Oxley v Regional Properties Ltd* [1944] 2 All ER 510, 512.

58 *Copley v Newmark* [1950] VLR 17, 19.

59 Lloyd Jones J in *Brumwell v Powys County Council* (n 55) [28].

60 Geoffrey Lane LJ in *Aldrington Garages Ltd v Fielder* (1983) 7 HLR 51, 60.

61 *AG Securities v Vaughan; Antoniades v Villiers* (n 31).

62 Bridge (n 24) 346.

is not to say that they do not equally appropriately find their place in a licence'.[63] A cosmetic difference, however, is that the licence will be drafted in language redolent of a non-proprietary relationship. The agreement will be labelled as a licence, identify the parties accordingly, avoid any reference to the word 'rent' and expressly deny exclusive possession. Against this backcloth, the role of the court is to construe the contractual documentation in an even-handed manner and divine a meaning corresponding to 'the intention of the parties, objectively ascertained by reference to the language and relevant background'.[64] Of course, the court is well aware that, even if both parties have signed the contract, this does not mean that they intended to be bound by those terms. This is of heightened concern within the residential sector where avoidance measures have historically been rife.

While exclusive possession (i.e. the legal right to exclude others from the property)[65] had long been prized as the decisive indicator of a tenancy,[66] this was to change during the decades preceding the House of Lords' decision in *Street v Mountford*. A series of seminal cases featuring Lord Denning forged a novel approach, which relegated the importance of exclusive possession. In doing so, Lord Denning 'began the trend of making what had seemed clear and self-evident subtle and confusing'.[67] The majority of these decisions were reached in the context of the Rent Acts, but the Denning approach was to apply also to commercial properties.[68] Although a finding of exclusive possession still offered prima facie evidence of a tenancy, it was no longer to be conclusive in favour of a lease, 'if the circumstances negative any intention to create a tenancy'.[69] The intentions of the parties were, thereby, promoted as the relevant test[70] and this theme was quickly reinforced in *Cobb v Lane*.[71] Lord Denning MR advanced the cause further in *Abbeyfield (Harpenden) Society Ltd v Woods*, where he acknowledged that:

> The modern cases show that a man may be a licensee even though he has exclusive possession, even though the word 'rent' is used, and even though the word 'tenancy' is used. The court must look at the agreement as a whole and see whether a tenancy really was intended.[72]

Some 10 years later, and in self-congratulatory mode, Lord Denning MR admitted to having 'revolutionised' the law over the preceding 25 years.[73] He emphasised that the issue lying at the core of the lease or licence distinction was whether it was 'intended that the occupier should have a stake in the room or did he have only permission for himself personally to occupy the room, whether under a contract or not?'[74] The techniques employed to discern

63 (n 11) 845.

64 Lord Hoffmann in *Bruton v London & Quadrant Housing Trust* [2001] 1 AC 406, 413.

65 As Lord Templeman commented in *Street v Mountford* (n 12) 816: 'A tenant armed with exclusive possession can keep out strangers and keep out the landlord unless the landlord is exercising limited rights reserved to him by the tenancy agreement to enter and view and repair'.

66 See *Lynes v Snaith* [1899] 1 QB 487; *Glenwood Lumber Co v Phillips* [1904] AC 405.

67 S Tromans, 'Leases and Licences in the Lords' (1985) CLJ 351, 352

68 See *Matchams Park (Holdings) Ltd v Dommett* (1984) 272 EG 549 (CA) (sports stadium); *Manchester CC v NCP* (1982) 262 EG 1297(CA) (car park).

69 Denning LJ in *Errington v Errington* [1952] 1 KB 290, 297. He added, at 298, 'if the circumstances and the conduct of the parties show that all that was intended was that the occupier should be granted a personal privilege, with no interest in the land, he will be held to be a licensee only'.

70 A change which Street (n 33) 328 described as marking 'a shift of emphasis from status to contract'.

71 [1952] 1 All ER 1199. It was also followed by the lower courts, for example, in *Murray Bull & Co Ltd v Murray* [1953] 1 QB 211, the court did not even debate whether exclusive possession had been granted.

72 [1968] 1 WLR 374, 375.

73 *Marchant v Charters* [1977] 1 WLR 1181, 1184.

74 Ibid 1185.

the 'intentions of the parties' were, however, cumbersome and not always in harmony with factual reality. An inherent irony was identified by Lewison: 'What was once a consequence of a finding that a licence existed had been transformed into a determining factor in establishing the existence itself'.[75]

Although the term 'intention' appears as a simple word of well-understood meaning, it has generated much judicial consideration in the context of landlord and tenant law.[76] Some insight was provided by Lord Wilberforce, who claimed that: 'When one speaks of the intentions of the parties to the contract, one is speaking objectively . . . What the court must do must be to place itself in the same factual matrix as that in which the parties were'.[77] In doing so, the court must always discover what the parties intended to do and not what relationship they hoped to bring about.[78] The relevant matrix, however, is usually a construct of the written agreement between the parties. Hence, it was the outward expression of accord from which the parties' intentions were traditionally gleaned.[79] Accordingly, in an attempt to identify the intentions of the parties, the courts were regularly seduced by a balance sheet-style approach, segregating those terms that indicated a tenancy into the credit column and those consistent with a licence into the debit column.[80] As Tromans noted: 'If the debits outweighed the credits, then the occupant had merely permission to occupy as opposed to a stake in the property.'[81] In cases where the agreement accurately states the intentions of the parties, the matter is straightforward 'Unless the parties' professed intentions differed from their true intention, or failed to reflect the true substance of the real transaction, this is conclusive'.[82] If, of course, the parties misunderstand the nature of their agreement, the court must identify and enforce the real transaction. This will arise where, as Mustill LJ explained, 'the language of the document (and in particular its title or description) superficially indicates that it falls into one legal category, whereas when properly analysed in the light of the surrounding circumstances it can be seen to fall into another'.[83] In both scenarios, the court is simply exercising its skills of construction so as to give fair effect to the true agreement on the terms expressed therein.[84]

Different considerations, however, emerge when it is alleged that the agreement contains an artificial provision inserted for the purpose of negating statutory protection or is in its overall nature a sham transaction. The court is then invited to adopt a broader approach and look beyond the written agreement to investigate the substance and reality of the transaction. The task is one which is undertaken with regard to such surrounding circumstances as the size of the property, the nature of the relationship between the parties, any pre-contractual negotiations and the conduct of the parties. The artificiality might be readily apparent in the context of the agreement, neon-lit when projected against the

75 K Lewison, *Lease or Licence: The Law after Street v Mountford* (Longman 1985) 5.

76 See M Haley, 'Section 30(1)(g) of the Landlord and Tenant Act 1954: The Unjust Relegation of Renewal Rights' (2012) 71(1) CLJ 118.

77 *Reardon Smith Line Ltd v Hansen-Tangen* [1976] 1 WLR 989, 996, 997.

78 *Cameron v Rolls-Royce* (n 36).

79 See *Eastleigh BC v Walsh* [1985] 1 WLR 525.

80 As, indeed, occurred in the Court of Appeal in *Street v Mountford* (1984) 16 HLR 27.

81 Tromans (n 67) 352. He added, ibid: 'Exclusive possession was merely one item on the balance sheet to be thrown in on the credit side: a weighty item no doubt, but only one item nonetheless.'

82 Millett J in *Camden LBC v Shortlife Community Housing Ltd* (1993) 25 HLR 330, 345.

83 (n 60) 1019.

84 As Geoffrey Lane LJ remarked in *Aldrington Garages Ltd v Fielder* (n 60) 62: 'If what the parties have agreed is truly a licence and not a tenancy dressed up in the verbiage or trappings or clothing of a licence then the owner is entitled to succeed.'

known factual matrix.[85] Such would occur when, say, exclusive possession is denied by use of a sharing clause, but there exists no space physically to accommodate another.[86] Similarly, the sharing clause might be so unlimited that it is clearly fictitious for, as Lord Jauncey observed of such a provision: 'If the clause is read literally the licensor could permit any number of persons to share the flat with the two defendants, even to the extent of sharing the joys of the double bed'.[87]

In less obvious cases, however, there are three major obstacles to a broad-ranging inquiry. The first is the lawyer's inbred hesitancy (absent a claim for rectification based on mistake) to disregard the unequivocal language of the agreement.[88] As Lord Goddard CJ observed: 'Written agreements, more especially when executed after professional advice and assistance are not to be lightly set aside.'[89] Until *Street v Mountford*, the fact that the overriding purpose was to avoid legislative controls did not itself invite scrutiny and the judicial attitude towards the circumvention of regulatory controls was one of complacency. As Lewison acknowledged, 'by a carefully drafted agreement whose effect was clearly understood by all parties a property owner could effectively contract out of the Rent Act'.[90] This view was shared also by Griffiths LJ who bemoaned 'how easy it is for a landlord to avoid the provisions of the Rent Act'.[91] The most notorious illustration of this lack of oversight was offered in *Somma v Hazlehurst*, where Lord Denning MR adopted a formalist approach and, by upholding a patently unrealistic sharing clause, gave effect to a licence agreement.[92] This decision was so shocking that it was even denounced in Parliament.[93]

The second obstacle takes the form of the parol evidence rule, which prevents a party to a written contract from adducing extrinsic evidence that contradicts, varies or qualifies the written terms of the contract. This rule of construction can be overcome only in strictly delimited circumstances. There has to be something about the agreement that invites the court's scrutiny, for example, an absence of express declaration of purpose, a lack of clarity on the face of the contract as to its purpose or some doubt cast on the veracity of the contract.[94] This could arise where the contract is too elaborate with regard to the transaction, much of the documentation is redundant or there is an important inconsistency in terms.[95] At this point, the court can take into account any relevant, extraneous facts within the knowledge of both parties and employ this information to give appropriate effect to the agreement.[96] Such extrinsic considerations have included, for example, the

85 In *Aslan v Murphy* [1989] 3 All ER 130, the agreement excluded the occupier from the accommodation for 90 minutes each day as a ploy to deny him exclusive possession. The Court of Appeal held that this term was colourable and utterly unrealistic.

86 *Demuren v Seal Estates* (1983) 7 HLR 83.

87 *AG Securities v Vaughan; Antoniades v Villiers* (n 31) 476.

88 See generally P Sparkes, 'Breaking Flat Sharing Licences' (n 13).

89 *R v Fulham, Hammersmith and Kensington Rent Tribunal ex p Zerek* [1951] 2 KB 1, 7. This embraces the manner in which the agreement describes itself: *Horford Investments Ltd v Lambert* [1974] 1 All ER 131.

90 Lewison (n 75) 11.

91 *Street v Mountford* (1984) 16 HLR 27 (CA), 40.

92 [1978] 2 All ER 1011. A similarly dubious clause was also upheld in *Aldrington Garages* (n 60).

93 Rt Hon Mr Frank Allaun, HC Deb 6 December 1978, vol 959, col 1403, declared it to be an 'astonishing case' and 'an abuse of the [Rent] Acts'.

94 *Samrose Properties v Gibbard* [1958] 1 All ER 502.

95 See *Walsh v Griffiths-Jones* [1978] 2 All ER 1002.

96 *Addiscombe Garden Estates Ltd v Crabbe* [1958] 1 QB 513. There a licence agreement to occupy a lawn tennis club was rejected.

intelligence and education of the parties,[97] the full appreciation by the parties of an agreement freely entered into[98] and the fact that both parties were businessmen.[99] Importantly, the parol evidence rule does not apply when there is an allegation of sham.[100]

The third obstacle concerns the accepted, but restrictive, definition of a sham offered in *Snook v London & West Riding Investments Ltd*.[101] There, Diplock LJ established the seemingly indispensable requirement that the parties 'must have a common intention that acts or documents are not to create the legal rights and obligations which they give the appearance of creating'.[102] Accordingly, a sham arises where the parties say one thing while intending another. The inhibiting nature of this rule of common intention was evident from the Court of Appeal decision in *Antoniades v Villiers*.[103] Although the two agreements in issue were the product of considerable artifice, they were not dismissed as sham agreements. The appellate court concluded that it was necessary for all the parties to intend that the occupiers had exclusive possession. As the landlord plainly did not share that intention, there could be no sham. Although the House of Lords was to reverse this decision,[104] it does reveal the traditional strictures associated with the identification of sham transactions. Indeed, Lord Templeman later expressed the wish that he had employed the word 'pretence' in his speech in *Street* rather than the phrases 'sham devices' and 'artificial transactions'.[105]

Accordingly, it is not enough that there exists an ulterior motive or that the agreement overtly benefits the licensor.[106] Similarly, it matters not that the court disapproves of the transaction.[107] If there is nothing done to disguise the substance of the agreement, the court cannot interfere.[108] The tenants in *Demuren v Seal Estates*, however, were able to demonstrate a sham because there was evidence of a prior oral agreement for a lease.[109] In *Grimsby College Enterprises Ltd v Revenue & Customs Commissioners*, Briggs J would have no truck with a licence agreement, which he regarded as 'an artifice, designed to present to the outside world in general (and, no doubt, HMRC in particular) a picture of the relationship between the parties very different from that which had been agreed'.[110] Nevertheless, it cannot be doubted that the onus of proving subterfuge and connivance rests heavily on the shoulders of the party contending that the written agreement should not be taken at face value.[111]

It is hardly surprising that, in the rental housing context, this narrow, judicial approach generated much contemporary criticism.[112] Clarke lamented the absence of 'a measure of

97 *Somma v Hazlehurst* (n 92).

98 *Sturolson & Co v Weniz* (1984) 272 EG 326. The absence of misrepresentation, mistake or undue influence is indicative of whether a party intended to be bound by the contract as drafted.

99 *Mackenzie v Kennett* (1962) 181 EG 9.

100 *Elmdene Estates Ltd v White* [1960] AC 528.

101 [1967] 2 QB 786.

102 Ibid 802.

103 (1988) 56 P&CR 334.

104 [1990] 1 AC 417. It held that the two contracts were in reality one and, so construed, the sharing clause was, as Lord Bridge admitted (at 454), 'repugnant to the true purpose of the agreement'.

105 Ibid 462; see S Bright, 'Beyond Sham and into Pretence' (1991) 11 OJLS 136.

106 *Farrell v Alexander* [1977] AC 59.

107 As Adams J stated in *Donald v Baldwyn* [1953] NZLR 313, 321: 'The law will not be hoodwinked by shams: but real and lawful intentions cannot be dismissed merely because they are disliked.'

108 *Brumwell v Powys County Council* (n 55).

109 (n 86); see also *O'Malley v Seymour* (n 56).

110 (n 14) [22].

111 See *Hadjiloucas v Crean* (n 83).

112 See, for example, A Arden, 'High Court Guerillas' (May 1979) Roof 78.

sanity and logic to this area of the law'.[113] Partington criticised the courts for having adopted a piecemeal approach, incrementally built up on a case-by-case basis, rather than devising general principles and concluded that: 'The result of these cases is to leave the law in turmoil.'[114] Gray and Symes regretted 'a more general retrenching of judicial opinion in favour of the entrepreneurial interest as opposed to the residential tenant'.[115] Griffiths LJ felt that the distinction between a lease and licence had become so eroded that: 'Parliament may have to consider bringing licences under the same umbrella of protection as tenancies.'[116] The legislature itself believed that the law had been brought into ridicule and disrepute.[117] It is against this background of academic, judicial and political dissatisfaction that Lord Templeman's speech in *Street v Mountford* is to be evaluated and the rejection of the licence as drafted by Mr Street understood.[118]

The legacy of *Street*

The setting for *Street v Mountford* was straightforward: it concerned residential property and the circumvention of the Rent Act 1977; there was no allegation of a sham agreement; it was not a sharing case and exclusive possession was not denied.[119] It was, therefore, a simple matter of construing an agreement that was styled as a personal, non-assignable licence in circumstances where direct contracting-out of the governing legislation was impermissible. At first instance, Ms Mountford was held to be a tenant, but this finding was overturned in the Court of Appeal.[120] As Slade LJ acknowledged: 'I do not see how the plaintiff could have made much clearer his intention that what was being offered to the defendant was a mere licence to occupy and not an interest in the premises as tenant'.[121] As there was no misrepresentation as to the parties' intentions, this outcome was on all fours with the Denning authorities. The House of Lords, however, unanimously reversed this decision and, in doing so, jettisoned decades of entrenched, albeit lax, legal reasoning as authored by Lord Denning. This, as Mr Street subsequently complained:

> turned the clock back more than a quarter of a century . . . The ancient wisdom is reinstated . . . The parties' intention as to the interest they wished to create is entirely irrelevant, however genuine and however unambiguously expressed.[122]

The obvious difficulty faced by Lord Templeman was, having disapproved of many of the pre-existing authorities, how in future the court was to distinguish between a tenancy and a residential licence. The House was, as Partington explained, 'torn between a desire to uphold contractual arrangements and an awareness, however dimly perceived, that the Rent

113 D Clarke, '*Street v Mountford*: The Question of Intention – A View from Down Under' (1986) Conv 39, 43.

114 Partington (n 54) 337. With acute foresight he added, ibid: 'Unless one of these cases reaches the House of Lords, and is heard by a particularly robust court, the chances of the judiciary providing a broader set of principles seem remote'.

115 K Gray and P Symes, *Real Property and Real People* (Butterworths 1981) 424. The authors' lament was that recent decisions had 'inflicted serious harm upon the social philosophy expressed in the Rent Act' (ibid).

116 *Street v Mountford* (n 80) 40. An idea mooted also in the Law Commission report (n 34) [4.6].

117 Rt Hon Mr Donald Anderson observed that such narrow legalism 'can only add substance to the claims of those who say that the courts are wholly out of touch with the needs of ordinary people' (HC Deb 7 March 1979, vol 963, col 1312).

118 Mr Street later described himself as, 'a double rogue, a landlord and a lawyer' (n 33) 328.

119 Hence, as Mustill LJ pointed out in *Hadjiloucas v Crean* (n 83) 1022, 'the decision has nothing directly to say about the manner in which the existence of such an intention [to grant exclusive possession] should be ascertained'.

120 *Street v Mountford* (n 80).

121 Ibid 39.

122 Street (n 33) 329.

Acts have a social purpose that should not be too easily thwarted'.[123] Lord Templeman's policy-led solution was disarming in its straightforwardness in that 'the only intention which is relevant is the intention demonstrated by the agreement to grant exclusive possession for a term at a rent'.[124] The House, thereby, breathed new life into the orthodox notion that exclusive possession is, as a general rule, the crucial factor in determining the existence of a tenancy. As Lord Templeman acknowledged: 'No other test . . . appears to be understandable or workable.'[125] Few would object to this reassertion of exclusive possession as the dominant characteristic of a tenancy.[126] From this perspective, it matters not whether the tenancy is of residential, agricultural or commercial property. As it is to do with legal rights, rather than de facto use,[127] it sits well with traditional notions of private property and the vital ability to exclude all others from the land.[128] This is in contradistinction to a lodger who, being a character that features prominently in Lord Templeman's speech, 'is entitled to live in the premises but cannot call the place his own'.[129]

Although there can be no tenancy without exclusive possession, it is trite law that, as Lord Templeman explained, 'an occupier who enjoys exclusive possession is not necessarily a tenant'.[130] By way of an exception to the general rule, one scenario which has emerged in the commercial sphere concerns a prospective tenant who is allowed to occupy, pending the grant of a lease or sale. Although the occupier may enjoy exclusive possession, if the transaction never goes ahead the occupier will be a licensee during the interim period.[131] For this exception to operate, the parties must genuinely intend that the transaction will eventually take place.[132] As Mann J emphasised in *Cameron Ltd v Rolls-Royce plc*, the relevant transactions must be contractually interrelated and the licence must be 'ancillary to a relationship which is part of a bigger picture'.[133]

Lord Templeman was not, however, content with the re-establishment of exclusive possession as the distinguishing feature of a tenancy. He advocated that the judiciary should adopt an interventionist stance, proactively seeking out rogue terms and colourable agreements.[134] Within the parameters of the case before him, this invitation appears as understandable as it is reasonable. Directly within his line of sight was the private-sector residential landlord who sought indirectly to contract out of socially imperative

123 Partington (n 54) 336. As Lord Templeman later acknowledged in *AG Securities v Vaughan; Antoniades v Villiers* (n 31) 458: 'Parties to an agreement cannot contract out of the Rent Acts; if they were able to do so the Acts would be a dead letter.'

124 *Street v Mountford* (n 12) 826.

125 Ibid 824. He reinforced this fundamental proposition in *AG Securities v Vaughan; Antoniades v Villiers* (n 31).

126 But see Street (n 33) 328: 'Regrettably the apple cart has now been upset, and perhaps damaged beyond repair.'

127 An occupier can have the sole occupation of premises without having exclusive possession: *Vesely v Levy* [2007] EWCA Civ 367.

128 See generally K Gray, 'Property in Thin Air' [1991] CLJ 252.

129 *Street v Mountford* (n 12) 818. Albeit a quaint notion, a lodger is an occupier whose landlord provides attendances and services, which require unrestricted use and access to the premises.

130 *Street v Mountford* (n 12) 818. He provided the examples of when there was no intention to create a legal relationship or the relationship between the parties was that of a vendor or purchaser, master and servant occupier, or similar.

131 *Isaac v Hotel de Paris Ltd* [1960] 1 WLR 239.

132 *Essex Plan v Broadminster* (n 37).

133 (n 36) [22].

134 See C Harpum, 'Leases, Licences, Sharing and Shams' (1989) CLJ 19.

legislation.[135] Outside this sphere, his call to action is both controversial and unhelpful. The use of licence agreements within the commercial context fell well outside his range of vision. This is a very different world, which is not populated with lodgers and in which it is possible directly to contract out of the statutory protection otherwise afforded. Three decades on from Lord Templeman's speech, the danger still remains that 'these passages are plucked from their context'.[136] It is incontrovertible that the existence of exclusive possession can readily be denied as regards business premises and the court must be particularly chary of interfering with a bargain voluntarily entered into between commercial parties. The factual matrix in which the agreement is sited must always be crucial and the extreme differences between the types of residential property and, say, a market stall, office or car park simply cannot be ignored.[137] The multiplicity of uses to which business premises may be put,[138] coupled with the diverse range of terms that may genuinely be agreed, serve only to underscore the fundamental differences that exist.[139] It necessarily follows that 'the indicia, which may make it more apparent in the case of a . . . residential occupier that he is indeed a tenant, may be less applicable or be less likely to have that effect in the case of business tenancies'.[140]

Although the skills of legal drafting remain prized, the unavoidable consequence of *Street v Mountford* is that the outward form of the agreement can no longer prevail over its substance. This entails that 'the categorisation of the agreement as a licence or a tenancy is for the law to decide, not the parties'[141] and, 'the parties cannot alter the effect of the agreement by insisting that they only created a licence'.[142] Lord Templeman was, thereby, signalling to the lower courts a preparedness to disregard artificial provisions that gave the misleading appearance that exclusive possession was denied. The court, as he later commented, 'must pay attention to the facts and surrounding circumstances and to what people do as well as to what people say'.[143] Somewhat optimistically, Lord Templeman believed that distinguishing between fact and fiction would be a straightforward task, adding that: 'In the case of residential accommodation there is no difficulty in deciding whether the grant confers exclusive possession'.[144] On this point, his judgment proved to be flawed and, as Hill concluded: 'Subsequent cases have shown that the formula suggested by Lord Templeman is in practice too simplistic'.[145] Even in the residential sector, difficulties were generated by the more sophisticated use of the sharing clause[146] and the employment of

135 As he admitted in *AG Securities v Vaughan; Antoniades v Villiers* (n 31) 459: 'The duty of the court is to enforce the Acts and in so doing to observe one principle which is inherent in the Acts and has been long recognised, the principle that parties cannot contract out of the Acts.'

136 Mustill LJ in *Hadjiloucas v Crean* (n 83) 1020.

137 See *Hunts Refuse Disposals Ltd v Norfolk Environmental Waste Services Ltd* [1997] 1 EGLR 16.

138 For example, the use of land and a farmyard with sheds for peat extraction and storage was upheld as a licence in *Crow v Waters* [2007] 2 P&CR DG14.

139 As Hutchison LJ explained in *Hunts Refuse Disposals* (n 137) 18, 'while one would ordinarily expect that someone in occupation of a small house for a fixed term at a rent had exclusive possession, one would I suggest have no such preconceptions about a person given the right to tip rubbish in the excavated parts of a large plot of land, on other parts of which, it seems, quarrying was continuing'.

140 Glidewell LJ in *Dresden Estates Ltd v Collinson* (1988) 55 P&CR 47, 52.

141 Mustill LJ in *Nicolaou v Pitt* (1989) 21 HLR 487, 492.

142 Lord Templeman in *Street v Mountford* (n 12) 819.

143 Lord Templeman in *AG Securities v Vaughan; Antoniades v Villiers* (n 31) 464.

144 *Street v Mountford* (n 12) 817.

145 J Hill, 'Shared Accommodation and Exclusive Possession' (1989) 52 MLR 408, 409.

146 See *AG Securities v Vaughan; Antoniades v Villiers* (n 31) .

so-called 'mobility clauses' (requiring the occupier to move to other accommodation if deemed necessary by the landlord).[147]

An attempted denial of exclusive possession in a private-sector residential licence agreement is, understandably, viewed with much caution and suspicion by the courts and is subject to the closest scrutiny. Post-*Street*, landlords who seek successfully to deny exclusive possession in that sector are driven to such elaborate lengths as providing services[148] or manufacturing a multi-occupancy scheme within which each occupier signs a distinct licence agreement containing different rights and obligations.[149] Matters are, however, much different in the commercial sphere where the licensor can genuinely reserve a high degree of control over the premises and the presence of exclusive possession is not readily to be inferred.[150] An authentic retention of territorial control, coupled with an express denial of exclusive possession, must ensure that the licence agreement is upheld. Accordingly, in *Shell-Mex & BP Ltd v Manchester Garages*,[151] a licence agreement prevailed in circumstances where the possession and control retained by the licensor (particularly concerning the products that could be sold and layout and equipment on the forecourt) ran contrary to the grant of exclusive possession. Similarly, in *Smith v Northside Developments*[152] no grant of exclusive possession could arise where there was a sharing of floor space between the occupiers, or in *Clore v Theatrical Properties Ltd*[153] where the licensee occupied under a trade concession agreement to sell refreshments in a theatre. Other contexts in which exclusive possession has been held not to exist include where occupation is limited to only part of the day or week;[154] where the right is to occupy part of the premises for storage and the licensor has the right to vary which part can be used;[155] and where there is occupation of a lock-up market stall.[156] This offers safe ground for judges who must necessarily conclude that there is no tenancy.

Problems arise when it is less evident that exclusive possession has been denied both at law and in fact. While the court might instinctively be repelled by any notion of 'the judge awarding points for drafting',[157] it is at this point that other indicators are sought. Indeed, there are a series of contractual terms that, whether by inclusion or omission, tend to signpost the true nature of the relationship. The covenant for quiet enjoyment, a reference to 'rent' and the reservation of a right of re-entry point towards a lease. In addition to an express denial of exclusive possession, the lack of a covenant for quiet enjoyment and the absence of a forfeiture clause suggest a licence agreement.[158] A blanket prohibition on alienation is also indicative of a licence agreement.[159] This allows the court to take an impressionistic overview of the transaction. Such occurred in *Venus Investments Ltd v Stocktop*

147 Glidewell LJ explained in *Dresden Estates* (n 140) 53: 'You cannot have a tenancy granting exclusive possession of particular premises, subject to a provision that the landlord can require the tenant to move to somewhere else'.

148 *Huwyler v Ruddy* (1996) 28 HLR 550.

149 *AG Securities v Vaughan; Antoniades v Villiers* (n 31).

150 *Venus Investments Ltd v Stocktop Ltd* [1996] EGCS 173 (CA).

151 (n 11) 841.

152 *Smith v Northside Development* (n 40).

153 [1936] 3 All ER 483.

154 In *Manchester City Council v NCP Ltd* [1982] 1 EGLR 94, the Court of Appeal saw sound commercial sense in such an arrangement.

155 *Dresden Estates v Collinson* (n 140).

156 *Gloucester City Council v Williams* (1991) 155 LG Rev 348.

157 Lord Templeman in *Street v Mountford* (n 12) 826.

158 *Shell-Mex & BP* (n 11); *Esso Petroleum v Fumegrange* (n 42).

159 *Barnes v Barratt* [1970] 2 QB 657.

Ltd,[160] where the agreement contained neither a forfeiture clause nor a right for the grantor to enter and inspect the premises. As it did not resemble a commercial lease, the Court of Appeal held it to be a licence agreement. A similar broad-brush approach was adopted in *NCP Ltd v Trinity Development Co (Banbury) Ltd*.[161] There, the Court of Appeal found it to be telling that the agreement, to occupy a multi-storey car park, was unusually framed so as only to impose a series of obligations on the occupier.[162] It was significant that the agreement did not start, as a tenancy would, with a conferral of an express right of occupation. Arden LJ emphasised that an absence of terms, characteristically found in a tenancy agreement, evidenced that a licence had been created. She felt able to uphold this highly stylised contract even though the degree of territorial control retained by the landowner was modest.[163] For those who advocate a strict adherence to the views of Lord Templeman, this will be viewed as a highly dubious decision. Such proponents would argue that an insufficiency of control must entail that the express denial of exclusive possession is a fiction. Although it is denied at law, it is not denied in fact. Hence, the conclusion drawn by Arden LJ might be regarded as the illegitimate promotion of form over substance. Nevertheless, the outcome was desirable and the means by which it was justified embodied both commercial realism and robust common sense.

Most contentious, however, is the weight to be given to the descriptive labels attached by the parties to their ostensible agreement. Although Lord Templeman protested that: 'Words alone do not suffice. Parties cannot turn a tenancy into a licence merely by calling it one',[164] in the commercial sector such descriptors are taken to offer prima facie evidence of the parties' intentions.[165] The difficulty lies in gauging the weight to be given to the professed intentions of the parties. In the residential rental market, scant attention is paid to such declarations. Since the *NCP* case, however, a very different ethos prevails in the commercial sector. There Arden LJ appeared to toe the traditional line by accepting that the labels attached to the agreement were not determinative. In homage to Lord Templeman, she admitted that the court must instead have regard to the substance and not merely to the appearance of that agreement. Nevertheless, she quickly veered off-message and, waylaid by the outward form of the contract before her and the context in which it arose, conceded:

> It would in my judgment be a strong thing for the law to disregard totally the parties' choice of wording and to do so would be inconsistent with the general principle of freedom of contract and the principle that documents should be interpreted as a whole.[166]

She did, however, acknowledge that the interpretative worth of an express declaration of intention would vary according to the respective circumstances of the parties. By way of a concession to the residential rental market, she cautioned that 'it must be approached with healthy scepticism, particularly, for instance, if the parties' bargaining positions are asymmetrical'.[167] It was, therefore, central to her judgment that this was an arm's length

160 (n 150).

161 (n 18).

162 For example, to operate a car park, to pay for security, to insure against specified risks, to account for profits, to allow the licensor to carry out certain repair works and to provide for 40 free car-parking spaces for the licensor's nominees.

163 See M Haley, 'Licences of Business Premises: Principle and Practicality' (2001) Conv 348.

164 *Street v Mountford* (n 12) 821.

165 An examination of the agreement as a whole necessarily includes some consideration of what the parties had indicated they intended to do: *Dresden Estates v Collinson* (n 140).

166 (n 18) [28]. Buxton LJ at [42] agreed that it would, indeed, be an 'extreme' response to exclude from the construction process those parts of the agreement which stated the intention of the parties.

167 Ibid [26].

agreement between two business parties who, having been legally advised, could be taken to have appreciated the implications of creating a licence instead of a tenancy. It is clear that Arden LJ could see no valid reason why the court should seek to overturn such commercial agreements. For the disciple of Lord Templeman, however, this may appear as legal heresy and a throwback to a bygone era when the 'intentions of the parties' dominated legal thinking.

This context-driven approach to construction later received the approval of Jonathan Parker LJ, who agreed that, in the commercial sphere, the form of the contract could not be ignored.[168] He explained:

> Where the contract so negotiated contains not merely a label but a clause which sets out in unequivocal terms the parties' intention as to its legal effect, I would in any event have taken some persuading that its true effect was directly contrary to that expressed intention.[169]

He felt it inappropriate that a party who agreed that 'the contract should not create a tenancy, should then invite the Court to conclude that it did'.[170] The clear message is that, provided the contract is between commercial parties, professionally drafted and not shown to contain misleading terms, it will be taken at face value and for what it says. Without doubt, within the commercial arena this sounds a retreat from *Street v Mountford* and resurrects a jurisprudence that, in the residential sector, has long been discredited. Nevertheless, this judicial response simply demonstrates that Lord Templeman's promulgation of the law does not work well outside the field of private sector housing.[171] As Mustill LJ observed sagely: 'It is a matter for regret that this important jurisdiction . . . should appear to be governed by rules which do not always yield a direct and unequivocal solution'.[172]

Conclusion

Undoubtedly, *Street v Mountford* is a seminal case on the landscape of landlord and tenant law. It has exerted a major impact on judicial thinking and the refinement of the lease or licence distinction both on a jurisprudential and a practical level.[173] The social and political context in which *Street* was decided, moreover, made it inevitable that prior legal thinking, which had been predicated on the intentions of the parties as divined solely from the terms of the agreement itself, would be abandoned. The interests of social justice were to prevail over a contractual framework constructed to favour the interests of the residential landlord.

The reinstatement by Lord Templeman of exclusive possession as the decisive feature of a tenancy has proved to be both sensible and workable. The emphasis upon the substance and reality of the transaction, as tested by an examination of the surrounding circumstances, adds a sense of reality, which had previously been absent. The decision in *Street*, and particularly the directive that the courts should actively seek out pretence agreements, however, can only be understood and properly applied in the context of the Rent Acts and the perceived need to shield the vulnerable from the unscrupulous. In both *Street* and *AG Securities v Vaughan*, Lord Templeman alluded repeatedly to the basic policy considerations underlying Rent Act protection and the

168 *Clear Channel v Manchester CC* (n 23).

169 Ibid [29].

170 Ibid [28].

171 As Ralph Gibson LJ put it in *Crancour v Da Silvaesa* (n 20) 211: 'The task of the court in applying the principles laid down in *Street v Mountford* is not that of applying the words of a statute'.

172 *Hadjiloucas v Crean* (n 82) 1025.

173 O'Connor LJ explained in *Brooker Settled Estates Ltd v Ayers* (1987) 282 EG 325, 326: '[Lord Templeman] sought to introduce some order into the law for the better administration of the law and guidance of the judges, particularly in the county court'.

blanket prohibition against the contracting-out of its provisions. The rallying cry for the judiciary to adopt an interventionist stance was not designed to pervade the commercial sector, where there are few concerns about inequality of bargaining power and in which it is perfectly permissible to contract out of the governing legislative code. It is also significant that a licence of commercial premises represents a transaction which benefits the occupier as opposed purely to promoting the interests of the landowner. There is simply no need for judicial vigilance in this very different world where both parties are commercial entities, are well equipped to obtain legal advice and able to contract freely on such terms as they can best negotiate.

Nevertheless, the courts have clearly struggled to adapt the decision of the House of Lords to the commercial context. The Law Commission rightly predicted that: 'It is not likely to have ended the battle; rather it settled a series of skirmishes and moved the front line'.[174] At one extreme, some judges have denied that there is any difference whatsoever between the two sectors, whereas at the other some have advocated a return to ascertaining the intentions of the parties from the text of the written agreement. Of these polar opposites, the latter is clearly correct. It is indisputable that the modern licence agreement has a vital role in facilitating the occupation of commercial premises. There should be no obstacle preventing the parties from creating a licence agreement of business premises if that is what they desire. All that should be required is for the parties to structure their relationship appropriately and genuinely intend to be bound by those terms. If the courts fail to recognise the obvious differences between residential licences and their commercial counterparts, it might be necessary for Parliament to take more drastic steps. This could result in the repeal of the Landlord and Tenant Act 1954, which is likely to have negative effects in the longer term for those who seek a tenancy agreement. Alternatively, the legislature could subject contractual licences to the same statutory regulation as tenancies and, thereby, 'eliminate the importance of the distinction between leases and licences in a large number of cases'.[175] As neither option is attractive or viable, it is to be hoped that the judiciary will, instead, adopt robust common sense to the construction of licences of business premises and recognise that, in the commercial world at least, rights of occupation should not depend on distinctions that are otherwise so difficult to draw.

174 Law Commission (n 34) [4.8].

175 Ibid [4.9].

NILQ 64(4): 443–63

On 'ring-fencing' the Common Foreign and Security Policy in the legal order of the European Union

Paul James Cardwell*

Reader, School of Law, University of Sheffield

1 Introduction

The European Union's (EU) 'unique' and 'supranational' legal order is widely regarded among lawyers to be crucial to its policy successes. If the EU is to become a successful global actor, one would expect its foreign policy and external relations[1] to be similarly built on supranational law. The legal aspects of the external relations of the EU were indeed at the heart of the debates on changes to the EU's constitutional structure in the Treaty of Lisbon 2007. But, unlike other policy areas, foreign policy was not brought within the 'supranational fold' by the Treaty text. On the contrary, the (then) UK Foreign Secretary characterised the position thus:

> Common foreign and security policy [CFSP] remains intergovernmental and in a separate treaty. Importantly . . . the European Court of Justice's [ECJ] jurisdiction over substantive CFSP policy is clearly and expressly excluded. As agreed at Maastricht, the ECJ will continue to monitor the boundary between CFSP and other EU external action, such as development assistance. But the Lisbon treaty considerably improves the existing position by making it clear that CFSP cannot be affected by other EU policies. *It ring-fences CFSP as a distinct, equal area of action.*[2]

As a statement by a UK politician to his national parliamentary chamber, this places a political spin on a significant legal development made by the Treaty of Lisbon to the Common Foreign and Security Policy (CFSP) and makes important assumptions about the development of the EU's constitutional order. Through considering the legal changes to the

* p.cardwell@sheffield.ac.uk. I would like to thank Maja Brkan, Tamara Hervey, Andrew Johnston, Peter Van Elsuwege and the anonymous reviewers for their helpful comments, and Simon Banks and Matthew Flintoff for valuable research assistance. A previous version of this article was presented at the Modern Law Review Seminar, 'EU External Relations Law and Policy in the Post-Lisbon Era', School of Law, University of Sheffield (January 2011) and at the Sussex European Institute, University of Sussex (November 2012).

1 A note on terminology: 'external relations' is the preferred term when speaking of the EU's relationships and policies beyond its borders. 'Foreign policy' is generally only applied to analysis of the CFSP. The Treaty of Lisbon introduces 'external action' as an umbrella term covering both external relations and foreign policy. In this article, references are deliberately made to external action, external relations and foreign policy as appropriate.

2 Secretary of State for Foreign and Commonwealth Affairs (David Miliband), HC Debs 20 February 2008, col 378. Emphasis added.

CFSP and subsequent practices, this article examines the extent to which the CFSP is 'ring-fenced' from other aspects of EU competences, which also cover external relations, and what this means for the legal dimensions of the EU's capacity to act beyond its borders. In doing so, the article revisits fundamental questions about the nature and function of law within the CFSP and, in turn, its place in the EU's constitutional order. The article critiques the ring-fencing metaphor and contends that it is only partly useful in explaining the role and place of the CFSP, since the foreign policy the EU has committed itself to forge is unlikely to rely only on the CFSP, but also the myriad of other competences under the Treaties. As there have been only a few instruments used post-Lisbon which rely on CFSP competences, current practice shows that the development of EU foreign policy largely occurs outside the formal scope of the CFSP. Hence, the CFSP continues to serve as a political arena for the Member States seeking to prevent EU action on an issue of vital (national) importance and to show that Member States retain control over foreign policy, by pointing to the CFSP's ring-fenced nature. Yet foreign policy co-operation does not end with the CFSP. Rather, the consequence is that the EU institutions find ways of putting external policies into action via an increasing set of legal instruments. The downside, at least for those proponents of a more obviously workable EU foreign policy, is that the CFSP is likely to remain characterised as a failure because, although it occupies the most obvious Treaty-based 'heart' of the EU's external relations, it is not the legal basis for practical policy making.

The article proceeds as follows: after setting out the CFSP's position in the EU's post-Lisbon legal order, the article critiques the extent to which the ring-fencing is borne out in the text of the Treaty. Analysis of the jurisdiction of the Court of Justice (CJEU) and the Treaty-based loyalty clause suggests that the fence is not as secure as it may seem. The article then considers whether the ring-fencing in practice stands up to scrutiny, with emphasis on Lisbon's institutional innovations to ensure coherence and consistency, and contends that the practice is even further removed from the impression given in the text of the Treaty. The article concludes that the consequences for the CFSP are that it will remain largely declaratory in nature and closer to a model of classic international law. If this means attempting to separate a policy area from the 'normal' methods of integration, then there are significant consequences for the future of EU law as we know it.

2 The CFSP in the EU's post-Lisbon legal order: reinforcing a paradox?

The CFSP would, to a casual observer unfamiliar with the complexity of the EU's workings, lie at the central core of the EU's external relations. Indeed, the Treaty on European Union (TEU) lays down the expansive provision that: 'The Union's competence in matters of common foreign and security policy shall cover all areas of foreign policy and all questions relating to the Union's security.'[3] The provision was introduced in the very first version of the TEU 1992 as a means to strengthen the EU's voice in international affairs to a level consummate with its growing economic weight. The CFSP codified informal practices and discussions on foreign policy affairs between the Member States dating back to the 1970s, but with grand statements in the Treaty about the EU's aims that it has found difficult to live up to.[4]

The revised provisions and post-Lisbon practice emphasise the EU co-ordinating its external competences in a more coherent fashion. Even if the scope of the CFSP has not changed drastically, the position of the CFSP (and the institutional competences in it)

3 Article 24(1) TEU.

4 The earliest and most widely cited statement relating to this is that of C Hill and his 'capability-expectations' gap: C Hill, 'The Capability-Expectations Gap or Conceptualising Europe's International Role' (1993) 31 Journal of Common Market Studies 305–28.

within the Treaty arrangements has been significantly altered. The CFSP is also listed as a separate Union competence in Article 2(4) of the Treaty on the Functioning of the European Union (TFEU) to distinguish it from other, 'general' competences. This 'otherness' of the CFSP within the constitutional order is expressed in the ring-fencing metaphor. But it does not explain why the CFSP should be exceptional within the EU's legal order. In one sense there is an obvious answer: the tradition of otherness of the CFSP and legal expression of the Member States' fear of the encroachment on their sovereignty if the Court of Justice was able to extend supranational EU legal principles to foreign policy. The Treaty seems to stem the 'Brusselsization' of the CFSP where 'the member states have in practice entered a slippery slope of integration with decision-making competence "creeping" to Brussels'[5] with the Court in Luxembourg filling in the gaps. But given that other areas have been 'communitarianised' in the most recent Treaty, is the ring-fence likely to prove effective in keeping the CFSP separate from the rest of the EU's legal order? If, as demonstrated below through a discussion of post-Lisbon practice, this is highly unlikely, what are the potential consequences for both the EU's foreign policy and its constitutional/legal order?

The CFSP embodies a deep paradox at its core, which has been exacerbated by Lisbon. The amendments point to a strong, value-led approach to external relations.[6] Institutional innovations, notably the EU diplomatic service and foreign minister in all but name (the European External Action Service (EEAS) and the High Representative for Foreign and Security Policy), underline the importance of foreign affairs by attempting to improve institutional and representative 'practical' capacities. And yet, the Treaty maintains the legal inadequacy of the instruments provided for in the Treaty in order to meet these aims and objectives. In this respect, the position of the CFSP in the constitutional order is the most obvious area where stated aims lack the legal structures to bring about effective supranational policies.

3 The legal technicalities of ring-fencing the CFSP: pre- and post-Lisbon provisions

The creation of the CFSP in the TEU in 1992 led to the characterisation of the EU as formed by three 'pillars'. As the second pillar, the CFSP was accorded alternative instruments and processes distinct from the familiar first pillar regulations and directives and the 'Community method'. Due to the lack of extensive role of the supranational institutions, it was characterised as a largely intergovernmental pillar and lacking in legal dynamics.[7] The TEU also limited the Court's jurisdiction by the former Article 46 TEU which did not list the CFSP provisions as being within the Court's powers. This provision has been strengthened by Lisbon which mentions the Court specifically in the articles devoted to the CFSP.[8]

Yet, despite its intergovernmental tag, the variety and multilevel institutional actors involved in the CFSP, its complexity and its unpredictability[9] led to an increasingly

5 H Sjursen, 'Not So Intergovernmental after All? On Democracy and Integration in European Foreign and Security Policy' (2011) 18 Journal of European Public Policy 1078, 1090.

6 Article 21(1) TEU.

7 M Cremona, 'The Union as a Global Actor: Roles, Models and Identity' (2004) 41 Common Market Law Review 553, 571.

8 Articles 24 TEU and 275 TFEU.

9 Hill has commented that 'one can never be sure where the next decision is going to come'; C Hill, 'Convergence, Divergence and Dialectics: National Foreign Policies and the CFSP' in J Zielonka (ed), *Paradoxes of European Foreign Policy* (Kluwer 1998) 43.

widespread view that the CFSP was engaged in a process of 'progressive supranationalism', making its distinction from other areas of EU integration less clear-cut.[10]

The pillar structure was abolished by Lisbon and replaced references to the 'Communities' by references to the 'Union'. By granting explicit legal personality to the *Union* rather than simply the *Communities*,[11] the Treaty gives the impression that the TEU had brought the two intergovernmental pillars within the framework of the EU and, hence, placed the CFSP on the same footing as other, more integrated areas, ending its 'otherness'. In reality, this merely removed the strange situation where the Union relied on the legal personality enjoyed by the Communities to conclude international agreements and join, for example, the World Trade Organization and formalised the existing consensus that the EU had already become an independent subject of international law.[12] The abolition of the pillar structure did not mean the end of 'intergovernmental' areas of EU policy.

A related innovation in the attempt to bring together the CFSP and other externally-focused competences was a new TEU section on the need for consistency and coherence.[13] Article 23 TEU makes the CFSP subject to new general provisions on the Union's external action.[14] The three Treaty articles which govern the entirety of the CFSP and non-CFSP dimensions to the Union's activities are wide in scope and give some indication to the values the EU holds dear, though only a few are aimed towards specific goals.[15]

The 'specific provisions' applicable to the CFSP show that the former second pillar has not been fully flattened. Rather, the specific provisions contribute to the ring-fencing away from other law and policy-making areas. The Treaty retains but rewords the previous Article 47 TEU (now Article 40 TEU), which states the CFSP specific provisions shall not affect the exercise of Union competences in Articles 3–6 TFEU, essentially the former Community competences which include common commercial policy,[16] development co-operation and humanitarian aid[17] and other areas which have an external dimension, including freedom, security and justice, environment and energy. The subtle, but important, change contained with the post-Lisbon Article 40 is that, whilst the CFSP may not affect other competences, the reverse is also now the case, i.e. the use of regulations or directives in areas where the EU enjoys exclusive or shared competence with the Member States may

10 R Gosalbo Bono, 'Some Reflections on the CFSP Legal Order' (2006) 43 Common Market Law Review 337, 349.

11 Article 47 TEU.

12 A Sari, 'Between Legalisation and Organisational Development: Explaining the Evolution of EU Competence in the Field of Foreign Policy' in P J Cardwell (ed), *EU External Relations Law and Policy in the Post-Lisbon Era* (Asser Press 2012) 77; N Tsagourias, 'EU Peacekeeping Operations: Legal and Theoretical Issues' in M Trybus and N D White (eds), *European Security Law* (Oxford University Press 2010) 121. For further analysis of the organisations to which the EU is a party, see M Emerson, R Balfour, T Corthaut, J Wouters, P M Kaczyński and T Renard, *Upgrading the EU's Role as Global Actor: Institutions, Law and the Restructuring of European Diplomacy* (Centre for European Policy Studies 2011) 22–36.

13 On coherence, see M Cremona, 'Coherence in European Union Foreign Relations Law' in P Koutrakos (ed), *European Foreign Policy: Legal and Political Perspectives* (Edward Elgar 2011); C Hillion, 'Tous Pour Un, Un Pour Tous! Coherence in the External Relations of the European Union' in M Cremona (ed), *Developments in EU External Relations Law* (Oxford University Press 2008); M Broberg, 'Don't Mess with the Missionary Man! On the Principle of Coherence, the Missionary Principle and the European Union's Development Policy' in Cardwell (n 12); and, within the context of the ENP, B Van Vooren, *EU External Relations Law and the European Neighbourhood Policy: A Paradigm for Coherence* (Routledge 2012).

14 Articles 21–46 TEU.

15 For example, in Article 21(2)(e) TEU: 'the integration of all countries into the world economy, including through the progressive abolition of restrictions on international trade'.

16 Article 3(1)(e) TFEU.

17 Article 4(4) TFEU.

not be used instead of the specific provisions of the CFSP. This provision is likely to result in inter-institutional disputes before the CJEU which will be required to ascertain where the centre of gravity of a measure lies.[18]

The CFSP is the only area which applies to all Member States where 'specific provisions' for decision-making apply.[19] Unanimous voting in the Council remains the basis of the decision-making process (Article 24(1) TEU) and, even when abstaining, a Member State 'may qualify its abstention by making a formal declaration' (Article 31(1) TEU). The CFSP is alone within the Treaty therefore in retaining the 'Luxembourg compromise'[20] and permitting any Member State to free itself from the obligation to apply a CFSP decision, even though that decision will bind the EU.[21] In an enlarged Union of 28 Member States, one need look no further than the retention of unanimity in CFSP decision-making to illustrate its otherness in the legal order and the associated difficulties in decision-making. In the report of Working Group VII on External Action under the European Convention, it was noted that:

> Some members . . . expressed the opinion that foreign policy issues were not adapted to decision making by voting since it would be difficult for a Member State to find itself in a minority position on an issue in which precisely its national interests were at stake. Some pointed out that QMV [qualified majority voting] in CFSP would also heighten third country awareness of internal EU disagreement, thus rendering CFSP less effective.[22]

It is quite a curious claim to make that outside actors may seek to take advantage of EU difference, since this could apply to any of the areas in which the EU operates, especially those which are 'sensitive' in terms of national sovereignty but which have nevertheless witnessed a growing amount of EU competence. Therefore, the choices in the Treaty reflect the 'super-sensitive' nature of foreign policy with a built-in safeguard mechanism to ensure that any decisions have been agreed by one and all.

The Treaty-based instruments were renamed by the Treaty of Lisbon[23] but remain deliberately separate from the more familiar instruments used elsewhere.[24] Common strategies, created in the Treaty of Amsterdam as a means of structuring EU action on areas of focus, remain unchanged, but in practice are hardly ever used. The EU has preferred to agree 'strategic' documents which do not rely on a specific legal basis, including the European Security Strategy (2003) and Stabilisation and Association Process for South-East

18 I am grateful to Peter Van Elsuwege for helpful discussions on this point.

19 Articles 136–38 TFEU are also defined as 'specific provisions' but only for those Member States whose currency is the euro: the specificity is due to their application to certain Member States only rather than the policy area itself, as for the CFSP.

20 R Schütze, *European Constitutional Law* (Cambridge University Press 2012) 207.

21 Piris notes that this provision has only been used once, by Cyprus in 2008 regarding the EU Rule of Law mission in Kosovo: J-C Piris, *The Future of Europe* (Cambridge University Press 2012) 77.

22 European Convention, *Final Report of Working Group VII on External Action* CONV 459/02 (2002) 20.

23 CFSP measures take the form of decisions which define 'actions to be undertaken' (formerly 'joint actions') and 'positions to be taken' (formerly 'common positions').

24 It is worth noting here that pre-Lisbon instruments retain their validity, as per Treaty on European Union, Protocol 36. In Case C-130/10 *European Parliament v Council of the European Union* (Judgment of the Court (Grand Chamber), 19 July 2012) the Court reinforced this point (at para 109) by stating that: 'the fact that the EU Treaty no longer provides for common positions but for decisions in matters relating to the CFSP does not have the effect of rendering non-existent those common positions adopted under the EU Treaty before the Treaty of Lisbon entered into force'.

Europe.[25] 'Actions to be undertaken' and 'positions to be taken' which are made on the basis of 'decisions of the European Council on the strategic interests and objectives of the Union' are adopted by qualified majority in the Council,[26] as an exception to the usual rule of unanimity.[27] For proponents of a less intergovernmental CFSP, this provision appeared to offer an opportunity for majority voting which could have developed into the 'norm' of CFSP decision-making. Article 32 TEU points to the possibility of a 'common approach' on CFSP matters which could therefore be used for a similar purpose. However, a Member State may block a decision taken by qualified majority, if it conflicts with 'important and stated reasons of national policy' (Article 23(2)). This is one of the clearest factors contributing to the ring-fenced nature of the CFSP since such a provision allowing a national veto despite majority voting taking place is found nowhere else in the Treaty.[28]

Significantly, no formal enforcement mechanisms are provided for in order to ensure Member State compliance. The lack of formal enforceability of the CFSP instruments leads some to conclude that they cannot be considered to be 'legal' at all.[29] Others have noted the 'lowest common denominator' character of the CFSP instruments as they seek to accommodate the divergent interests of all the Member States,[30] preventing even the type of enforceable minimum harmonisation found elsewhere in EU law.

The CFSP instruments are not within the scope of Article 288 TFEU and the effectiveness of CFSP measures in national courts has been debated since the entry into force of the TEU.[31] That is not to say that 'traditional' enforcement measures would necessarily be appropriate for use in the CFSP, as even the European Commission recognises,[32] but the lack of any enforcement mechanisms sets the provisions apart from the rest of the European integration process. To have no means of enforcing the provisions leaves a significant gap in the EU's legal order unless the measures taken are the type which do not lend themselves to enforceability; but this would sit uncomfortably with the wide scope of the Treaty provisions.

Taking the continuation of previous legal mechanisms surrounding the CFSP with newer initiatives emerging from Lisbon, at the formal level the text of the Treaty does

25 This is explored further in P J Cardwell, 'EuroMed, ENP and the Union for the Mediterranean: Overlapping Policy Frames in the EU's Governance of the Mediterranean' (2011) 49 Journal of Common Market Studies 219. See also G De Baere, *Constitutional Principles of EU External Relations* (Oxford University Press 2008) 115; S Hix, *The Political System of the European Union* (Palgrave Macmillan 2005) 391; P Koutrakos, *EU International Relations Law* (Hart Publishing 2006) 399; and S Keukeleire and J MacNaughtan, *The Foreign Policy of the European Union* (Palgrave Macmillan 2008) 154–56.

26 Article 31(2) TEU.

27 Article 31(1) TEU.

28 The closest provision is the 'emergency brake' found in Article 83(3) TEU concerning directives on criminal procedure pursued through Article 83(1)–(2), which allows for the ordinary legislative procedure (i.e. including QMV) to be suspended where a Member State has concerns that a draft directive 'would affect fundamental aspects of its criminal justice system'. However, this provision does not carry the same connotations as Article 23(2) TEU.

29 P Eeckhout, *External Relations of the European Union: Legal and Constitutional Foundations* (Oxford University Press 2004) 396.

30 J Zielonka, *Europe as Empire* (Oxford University Press 2007) 143.

31 Cf, for example, A Dashwood, 'The Law and Practice of CFSP Joint Actions' in M Cremona and B De Witte (eds), *EU Foreign Relations Law: Constitutional Fundamentals* (Hart Publishing 2008) and K Lenaerts and T Corthaut, 'Of Birds and Hedges: The Role of Primacy in Invoking Norms of EU Law' (2006) 31 European Law Review 287.

32 'While the "Union method" has been a great success in the internal market, it would simply not work for CFSP', M Šefčovič, Vice-President of the European Commission, 'The Treaty of Lisbon–One Year on Cooperation in a Mature Institutional Framework', speech delivered at the European University Institute, Florence, 12 November 2010 <http://europa.eu/rapid/press-release_SPEECH-10–647_en.pdf>.

indeed appear to ring-fence the CFSP away from mainstream EU law to a greater extent than was previously the case. However, three of the main innovations, or at least, more explicitly worded dispositions, require further analysis to discern whether the Treaty text does effectively ring-fence the CFSP.

4 Testing the ring-fence

i) Exclusion of the jurisdiction of the CJEU

The exclusion of the jurisdiction of the CJEU appears in the second paragraph of Article 24 TEU.[33] The previous, pre-Lisbon version of the TEU made no mention of the powers of the Court in the CFSP articles (Title V). Furthermore, the CFSP is further ring-fenced away from the reaches of the CJEU within the provisions dealing with powers of the CJEU, in Article 275 TFEU.[34]

Taken together, there appear to be two fences protecting the CFSP from judicial supervision. Not only do the new provisions exclude review of the substance of CFSP measures, but they also eliminate any supervision over *procedural* irregularity, since the jurisdiction is limited to monitoring the competence boundaries or the legality of 'restrictive measures'.[35] It would not seem possible that Article 263 TFEU could be used to mount a judicial review challenge to the way in which a CFSP decision, even one concerning 'restrictive measures' such as sanctions on an individual, was made. Neither (it seems) could the alternative judicial review process, via a preliminary reference from a national court, be used.[36] Human rights challenges cannot engage the Court with regards to CFSP measures, even though EU foreign policy has already given rise to cases in the European Court of Human Rights[37] and the Treaty foresees eventual EU adhesion to the European Convention.[38] Article 47 of the Charter of Fundamental Rights of the Union (on the right to an effective remedy) is difficult to square with the exclusion of the Court from CFSP matters, which could conceivably affect the legal rights of individual citizens.[39] In its 2012 judgment in *Parliament v Council*, the Court rejected an argument by the European Parliament that it would be contrary to EU law to adopt measures having a direct impact on the fundamental rights of individuals and groups which excluded the participation of the Parliament.[40] The Court stated that the Charter binds all institutions (and therefore also when institutions are acting under the CFSP) but did not elaborate on Article 47 specifically.[41]

33 'The Court of Justice of the European Union shall not have jurisdiction with respect to these provisions, with the exception of its jurisdiction to monitor compliance with Article 40 of this Treaty and to review the legality of certain decisions as provided for by the second paragraph of Article 275 TFEU.'

34 'The Court of Justice of the European Union shall not have jurisdiction with respect to the provisions relating to the common foreign and security policy nor with respect to acts adopted on the basis of those provisions.'

35 Article 215(3) TFEU.

36 Article 19(3)(b) TFEU gives the authority to the Court of Justice to give preliminary rulings 'on the interpretation of Union law or the validity of acts adopted by the institutions'. The Article does not mention the CFSP but says (in Article 19(3)) 'in accordance with the Treaties'. See also M Brkan, 'The Role of the European Court of Justice in the Field of Common Foreign and Security Policy after the Treaty of Lisbon: New Challenges for the Future' in Cardwell (n 12) 100.

37 M-G Garbagnati Ketvel, 'The Jurisdiction of the European Court of Justice in Respect of the Common Foreign and Security Policy' (2006) 55 ICLQ 77.

38 See further, C Eckes, 'EU Accession to the ECHR: Between Autonomy and Adaptation' (2013) 76 MLR 645; and T Lock, 'Accession of the EU to the ECHR' in D Ashiagbor, N Countouris and I Lianos (eds), *The European Union after the Treaty of Lisbon* (Cambridge University Press 2012) 109.

39 Brkan (n 36) 106.

40 Case C-130/10 *Parliament v Council* (n 24) para 83.

41 Ibid paras 83–84.

The Treaty does foresee an exception when individual rights are at stake in Article 275 TFEU, which now allows the Court to review decisions affecting rights of natural/legal persons (brought under Article 263 TFEU) but only in cases where restrictive measures are placed upon them.[42] In all other instances, even if an individual was able to satisfy the extremely high threshold of the standing requirements for a non-privileged applicant seeking judicial review of an act of the EU institutions, the Court would not have jurisdiction to hear the claim. The lack of jurisdiction to review CFSP measures characterises the policy firmly as an area of executive-led 'high politics' in which it is assumed that individual rights are unaffected.

Further, it is worth recalling that the CFSP covers *all* aspects of foreign policy and is not defined as merely a residual category. Comparisons with domestic systems of Member States also run into difficulty when bearing in mind the status of the Parliament, which has neither the legislative involvement in CFSP decisions, nor the ability to use its position to bring actions before the Court, even if dressed as procedural. Save for the exception in Article 275 TFEU relating to restrictive measures, the 'rule of law' which the Treaty attaches to both its own system, and the values it purports to promote beyond its borders, is thus diminished and replaced by 'a rule of the executive'.[43] Article 75 TFEU, which involves the ordinary legislative procedure, allows for sanctions against individuals, within the setting of combating terrorism within the area of freedom, security and justice.[44] Yet, under Article 215 TFEU, which is engaged when a decision has been taken under the CFSP, restrictive measures against third countries, or natural or legal persons, may be adopted by the Council acting alone. The Parliament's attempt to challenge the use of Article 215 TFEU as a basis for sanctions, by claiming that it would be contrary to EU law to adopt measures capable of impinging directly on fundamental rights, was recently rejected by the CJEU in *Parliament v Council* since it would make Article 215 TFEU redundant. The Court did underline the general obligation (as per Article 51(1) of the Charter and *Kadi*)[45] for all Union institutions to safeguard fundamental rights. Nonetheless, it is difficult to see from the decision in *Parliament v Council* what would be the case if the Court found that necessary safeguards on fundamental rights were not included in a measure taken under Article 215 TFEU.

The discussion thus far in this part suggests that the ring-fences around the CFSP appear relatively secure. However, whilst it might be challenging to envisage a situation where individual rights are affected by a CFSP measure other than restrictive measures, it should not be forgotten that the Court in *Kadi* found a solution to a complex legal conundrum which expressly confirmed that fundamental rights are protected by a 'constitutional guarantee stemming from the EC Treaty as an autonomous legal system which is not to be prejudiced by an international agreement'.[46] It is not too much of a conceptual stretch to consider that the Court could refer to general principles of law and the protection of fundamental rights to engage in a more substantive review of CFSP

42 It is worth noting here that since this provision is new, it was not used in the famous *Kadi and Al Barakaat* cases (Cases C-402 and 415/05 *Kadi and Al Barakaat International Foundation v Council and Commission* [2008] ECR I-6351) which were brought on the basis of the implementation powers via a Regulation pursuant to a CFSP common position (formerly Article 60 and 301 EC, now found in Article 215 TFEU).

43 Schütze (n 20); P Eeckhout, *External Relations of the European Union: Legal and Constitutional Foundations* (2nd edn, Oxford University Press 2011) 499.

44 Article 75 TFEU refers to the aims set out in Article 67 TFEU.

45 Cases C-402 and 415/05 *Kadi* and *Al Barakaat* (n 42) paras 82–83.

46 Ibid para 316. See further, P J Cardwell, D French and N D White, 'Kadi and Al Barakaat International Foundation' [2009] 58 ICLQ 229, 232.

decisions and the continuation (as noted before in *Kadi* case) of an 'implicit or indirect jurisdiction' over the CFSP.[47]

The ring-fencing of the CFSP has also cast some doubt on the principles established by the Court in pre-Lisbon CFSP case-law. Prior to Lisbon, the EC Treaty enjoyed a superior place in the hierarchy between the two Treaties since the former Article 47 TEU prevented the TEU from affecting anything within the Community's competences. This is no longer the case: revised Article 1 TEU states that the two Treaties shall have equal legal value. The TEU's original intent was to prevent intergovernmental pillars from having an effect on the more integrated, Community pillar: the CFSP can no longer be subservient to former Community competences. This casts some doubt on the continued validity of the Court's view in *ECOWAS*,[48] where the Court said that 'the Union cannot have recourse to a legal basis falling within the CFSP in order to adopt provisions which also fall within a competence conferred by the EC Treaty on the Community'.[49] The new Treaty arrangements appear to counter the approach the Court adopted in *ECOWAS* which required CFSP joint actions to be implemented not only by means of other CFSP decisions but also by Community decisions and thus forging a 'holistic' view of (external) competences. The Court rejected the UK's submission in *ECOWAS* that the CFSP provisions were entirely separate.[50] In the only post-Lisbon case to address this issue, Advocate-General (AG) Bot in *Parliament v Council* made an interesting and potentially far-reaching observation in referring to the competences outside the CFSP 'in which the European Union enjoys complete freedom'.[51] He also rejected the Council's view that, in the case of the EU competences on combating terrorism via restrictive measures, the deciding factor on competence is 'internal' v 'external'.[52] Craig's observation that '[T]he presumption that "normal" EU law should predominate is deeply ingrained in the judicial psyche and will not easily be shifted'[53] is implicit in the views of AG Bot which suggest that the CFSP only operates in defined areas, notwithstanding the wide definition of CFSP and the general obligation on EU institutions to act within their respective competences.

'Purely' political measures are difficult to isolate from the measures which can be taken under the competences of the EU. Therefore, it could be argued that any potential measure which covers interaction with the outside world could (or even should) be done under the CFSP, once again remembering that the CFSP covers 'all areas' of foreign policy.[54] The important word is therefore 'all'. Concluding a purely economic agreement with a third state implies that foreign policy choices have been made to engage with that country.[55] If this was the case before Lisbon, then it is even more pertinent now since there is no implication that the (former) Article 11 TEU is nonetheless subservient to the principle that the EC Treaty takes precedence (former Article 47 TEU). Faced with a situation where the Court, as it is permitted to do so under the Treaty, examines the legal basis of a measure to decide whether it belongs under the CFSP (which is hence not susceptible to judicial review) and a non-

47 D Curtin, *Executive Power of the European Union: Law, Practices, and the Living Constitution* (Oxford University Press 2009) 180–81.

48 Case C-91/05 *Commission v Council* (ECOWAS/Small Arms) [2008] ECR I-3651.

49 Curtin (n 47) 190.

50 Ibid.

51 Case C-130/10 *Parliament v Council* (n 24), Opinion of AG Bot (31 January 2012). The original French reads 'dans les domaines hors PESC dans lesquels l'Union a toute liberté' para 82.

52 Ibid para 75.

53 P Craig, *The Lisbon Treaty: Law, Politics, and Treaty Reform* (Oxford University Press 2010) 415–16.

54 Article 24(1) TEU.

55 Garbagnati Ketvel (n 37) 84. See also Case C-124/95 *Centro-Com* [1997] ECR I-81, Opinion of AG Jacobs, para 41.

CFSP measure (which would be susceptible) it would seem logical to assume that the Court would have a natural preference for a measure which can be justiciable. Its reasoning could be derived from the commitment to the rule of law or the preamble of the Treaty underlying the integration of the EU as the ultimate goal. Reference could also be made to the Member States doing all they could to meet the aims of the Union, as distinct from the loyalty clause. A close reading of Article 40 TEU would suggest that the parity of the two Treaties protects the *acquis communautaire* built up in the former first pillar which cannot be undone by relying on the CFSP as a legal basis since this would potentially breach Article 40 TEU.[56]

Although one should be wary of predicting how the Court will use its discretion to police the boundary between the CFSP and other policies, there is little doubt that it will have the opportunity to do so, especially in cases similar to the post-Lisbon case of *Parliament v Council*, and the assertion here is that the ring-fencing is unlikely to exclude the Court as it might appear from an isolated reading of the Treaty text.

ii) The loyalty obligation

The loyalty obligation is a distinct, but related, issue on the Court's position vis-à-vis the CFSP. The obligation is expressed generally (since Lisbon, 'the principle of sincere cooperation') in Article 4(3) TEU, which has formed a crucial part of the Court's reasoning in its development of the EU's supranational legal order.[57] The Treaty obliges Member States to comply with the Union's actions and support the CFSP 'actively and unreservedly in a spirit of loyalty and mutual solidarity'.[58] In addition, Article 32 TEU expanded the previous Article 16 TEU, the first paragraph of which requires a high level of consultation within the Council and 'mutual solidarity'. Article 28(2) TEU states that decisions taken 'commit the Member States in the positions they adopt and in the conduct of their activity', which can be considered a dimension of the obligation, albeit one which operates differently to enforceable Community-developed concepts.[59]

Member States have become used to the interpretations by the CJEU on the duty of loyalty, including areas which fall within the sphere of external relations, albeit not explicitly on the CFSP.[60] The CJEU has continued to develop, and extend, the doctrine to promote consistency of EU external relations and the EU's legal order.[61] As Neframi has argued, the duty has shown both the potential and dynamics across the non-CFSP aspects of EU external dimensions for Member States to be bound by the Court of Justice's interpretation of the Treaties.[62] The *AETR/ERTA* case[63] established long ago the doctrine of implied external powers, a decision of high constitutional significance for the development of the

56 Garbagnati Ketvel (n 37) 91.

57 See, inter alia, Case 96/91 *Commission v Netherlands* [1982] ECR 1791; Case 14/83 *Von Colson and Kamann* [1984] ECR 1891; Case C-106/89 *Marleasing SA v La Comercial Internacional de Alimentación SA* [1990] ECR I-4135; and Case C-213/89 *R v Secretary of State for Transport ex p Factortame* [1990] ECR I-2433.

58 Article 24(3) TEU.

59 P Van Elsuwege, 'EU External Action after the Collapse of the Pillar Structure: In Search of a New Balance between Delimitation and Consistency' (2010) 47 Common Market Law Review 987, 911.

60 Case C-246/07 *European Commission v Sweden* [2010] ECR 3317. For detailed analysis, see G De Baere, '"O, Where is Faith? O, Where is Loyalty?" Some Thoughts on the Duty of Loyal Cooperation and the Union's External Environmental Competences in the Light of the PFOS Case' (2011) 36 European Law Review 405.

61 For further, see Eeckhout (n 43) 59–64. For a discussion of a recent extension of the *AETR/ERTA* doctrine by the CJEU, Case C-45/07 *Commission v Greece* [2008] ECR I-701, see M Cremona, 'Extending the Reach of the AETR Principle' (2009) 34 European Law Review 754.

62 E Neframi, 'The Duty of Loyalty: Rethinking its Scope through its Application in the Field of EU External Relations' (2010) 47 Common Market Law Review 323, 324.

63 Case 22/70 *Commission v Council (AETR/ERTA)* [1971] ECR 263.

EU's external relations across the board and one which found its basis in the duty of loyalty.[64] If *AETR/ERTA* was the first landmark, then *Pupino*[65] is its natural successor in terms of using the loyalty principle as a springboard to ensuring legal effectiveness in the former third pillar. The Court stated that:

> It would be difficult for the Union to carry out its tasks effectively if the principle of loyal cooperation, requiring in particular that Member States take all appropriate measures, whether general or particular, to ensure fulfilment of their obligations under European Union law, were not also binding in the area of police and judicial cooperation in criminal matters.[66]

Hence, the Court 'found' a binding loyalty obligation on the basis of the need for effectiveness of EU law. It thus prompted much discussion about whether this reasoning could be applied/extended to the CFSP too.[67] The Court did not have an opportunity to do so, although, on a case involving external relations, the Court found that Sweden breached the (then) Article 10 EC by acting unilaterally when a 'concerted common strategy' existed at EU level and that, contrary to Sweden's and others' views, the duty is not limited in scope.[68] The Court stated that it did not matter whether the area of competence under question was exclusive or shared, the important factor is whether a 'situation is likely to compromise the principle of unity in the international representation of the Union and its Member States and weaken their negotiating power'.[69] This is suggestive of a *Pupino*-style analogy and demonstrates that external relations are not beyond the limits of the Court's view of the loyalty principle, but the question did not relate specifically to a CFSP measure and the refusal of a Member State to adhere to it. At the very least, however, the Court's decision in *Commission v Sweden* points to the obligation on the Member States to remain 'silent' on matters pertaining to foreign policy, even when the EU does not enjoy exclusive competence.[70]

Do the post-Lisbon provisions prevent a *Pupino*-style argument extending to the CFSP? The Italian and UK governments argued in *Pupino* that there was no loyalty clause which could be applied as Article 10 EC applied only in the *Community* pillar, but the CJEU found one on its own logic of interpreting the aims and scope of the Treaty and hence applying a definition which transcended the pillars. Reading across the *dicta* from the Court by analogy to the CFSP loyalty provision would, pre-Lisbon, have been a possibility, should the Court have had an opportunity. But, as examined in the previous section, the Treaty text *appears* to now prevent an opportunity for the CJEU to do so. But even if the ring-fence secures the CFSP from other policy areas on the substance, that is not necessarily the case for the loyalty provisions. Given that the loyalty provision applies across the EU's legal order, loyalty cannot remain outside the CFSP. The problem for the Court, should it adopt this position, is the lack of jurisdiction: in other words, the Court would be able to extend the same characteristics of loyalty as it found in *Pupino* to the CFSP, but would need to circumvent the jurisdictional problem.

The most conceivable way the Court might be able to take such a step would be to return to its boundary-policing function by resorting, as it did in *Pupino*, to 'a (thin) textual

64 This doctrine is now contained within Article 3(2) TFEU.

65 Case C-105/03 *Criminal Proceedings against Maria Pupino* [2005] ECR I-5285.

66 Ibid para 42.

67 Gosalbo Bono (n 10) 366.

68 Case C-246/07 *European Commission v Sweden* [2010] ECR I-3317, para 70.

69 Ibid para 104.

70 A Delgado Casteleiro and J Larik, 'The Duty to Remain Silent: Limitless Loyalty in EU External Relations?' (2011) 36 European Law Review 524, 540–41.

argument as well as a more persuasive teleological one'.[71] In this context, the Court could argue that it was not the *substance* of a CFSP measure but the *competence* at stake, and that (returning to the lack of hierarchy between the Treaties) that loyalty requires the more integrationist legal basis to be used. This scenario is most likely to arise in a situation where the Parliament claims that the CFSP provisions are being improperly used. However, it is possible that, especially in a case where individual rights are affected, a case could arise via the preliminary reference procedure from a national court. This prompts a further potential difficulty, however, since national courts would not be able to seek clarification on how to resolve a conflict between national law and a CFSP measure via this procedure and would therefore have to come to their own conclusions as to how to interpret the non-legislative character of the measure which nevertheless is said in the Treaty to be subject to a loyalty clause. This is similar to Schütze's characterisation of external competence restraint as a form of 'reverse' subsidiarity.[72]

Much here depends on whether situations arise where there is a doubt over competence, as in *ECOWAS*. The Parliament can be expected to be particularly vigilant in this regard and has already begun proceedings against a CFSP decision on an EU agreement with Mauritius on combating piracy, which the Parliament feels does not fall wholly within the CFSP.[73] As the guardian of the Treaties, the European Commission could be expected to also be vigilant but is in a more difficult position since the High Representative is a Vice-President of the Commission.

Further, the loyalty clause in *Pupino* was 'read across' from the EC Treaty (Article 10 EC): as Article 4(3) is now found in the TEU and, bearing in mind the collapse of the 'pillar' system, there is a strong case to suggest that it applies more directly to the CFSP.[74] In this context, therefore, even if not in others, the argument that the Lisbon text ring-fences CFSP away from the loyalty provision is a difficult one to make, especially if the CFSP provisions are used to get around other provisions where loyalty is much more firmly established.

iii) The non-legislative character of the instruments

The Treaty is now explicit that the instruments provided for under the CFSP cannot be considered 'legislative'.[75] CFSP decisions are excluded from the Article 289(3) TFEU procedure, 'Legal acts adopted by legislative procedure shall constitute legislative acts', which means that Article 24 TEU leaves open the question of what legal effect(s) CFSP decisions may have. A distinction must be made between 'legislative' and 'legal' since the former refers to the process of creating an instrument rather than its (legal) force. Article 289(3) does not refer to legal acts which might be adopted *outside* the scope of the legislative procedure.

This question of legal effects of the CFSP is not new, since the original TEU provisions were no clearer, only that the instruments were distinct from regulations, directives and

71 E Spaventa, 'Opening Pandora's Box: Some Reflections on the Constitutional Effects of the Decision in *Pupino*' (2007) 3 European Constitutional Law Review 5, 22.

72 Schütze (n 20) 217. This reference was mainly directed at the mixed agreements of the Union, but seems equally applicable to the CFSP if we are to take it as having a binding quality.

73 Case C-658/11 *European Parliament v Council of the European Union* (decision pending).

74 C Eckes, 'EU Counter-Terrorist Sanctions against Individuals: Problems and Perils' (2012) 17 European Foreign Affairs Review 113, 129.

75 Article 24 TEU.

decisions.[76] If the distinction between measures taken under the CFSP and the former first pillar was the supposed political, intergovernmental nature of the former, which do not lend themselves to the type of instruments that should be enforceable, then this was surely buried by the *Kadi* decision.

Whilst excluding the possibility of creating legislative acts, this should not be taken to mean 'non-binding' or 'non-legal',[77] particularly since they may indeed affect the legal rights of natural or legal persons.[78] The best characterisation remains Curtin's pre-Lisbon analysis of CFSP instruments as examples of 'binding non-legislation'.[79] After all, the Treaty does refer to the binding nature of the CFSP and, referring back to the *Pupino* decision, it was this type of provision that led the Court of Justice to find that the obligation could be read across. The lack of explicit mention of the principles of direct effect and supremacy/primacy in the Treaty and to which aspects of EU law it applies, does not help discover whether there is a possibility of these principles applying to the CFSP.

If the Court of Justice could find that, despite the wording of Article 24 TEU, the characteristics of a decision taken under the CFSP point to the same legal effects, then this is potentially significant in the choice of legal basis for measures. The *ECOWAS* case[80] demonstrated the extent to which the choice of legal basis for a particular measure is important, but complex. This is especially the case when there are multiple objectives pursued in a measure and the 'high politics' cannot be easily separated from the economic.[81]

Since the Treaties now have equal weight, this analysis leads back to the same argument about how the CFSP might be used in the future. If it is used only when a measure cannot be taken under another, non-CFSP legal basis then the legal situation will remain the status quo. But, if the non-legislative CFSP decisions begin to be used instead, then the Parliament in particular will make its voice heard through the judicial review procedure. Although unsuccessful, the Parliament has already demonstrated its will in this regard in its action against the Council, and invoked the democratic principles cited in the Treaty.[82] The Court would then be faced with a choice of putting aside its previous jurisprudence and what the Treaty used to say about the Treaty hierarchy, or find some inventive way of keeping the CFSP free of measures which could/should be found elsewhere. In short, the Court could conceivably be called upon to judge whether a decision taken under the CFSP should rather be a decision under non-CFSP Article 288 TFEU.[83] To do so would therefore circumvent the apparent lack of legislative qualities of CFSP instruments.

5 Post-Lisbon practice in the CFSP

The obvious starting point in examining post-Lisbon practice is the use of the instruments, which have been used very sparingly since the entry into force of Lisbon. The majority of

76 One might also note here that the characteristics of directives were developed by the Court, rather than the text of the Treaty, and there has been vigorous debate about whether the Court was justified in doing so: see T Hartley, 'The European Court, Judicial Objectivity and the Constitution of the European Union' (1996) 112 Law Quarterly Review 95; and A Arnull, 'The European Court and Judicial Objectivity: A Reply to Professor Hartley' (1996) 112 Law Quarterly Review 411.

77 Van Elsuwege (n 59) 1005.

78 Article 215(2) TFEU. But see also Article 75 TFEU and Eckes (n 74) 120.

79 Curtin (n 47) 124.

80 Case C-91/05 C*ommission v Council (ECOWAS/Small Arms)* [2008] ECR I-3651.

81 P Koutrakos, 'Current Developments: European Union External Relations' (2010) 59 International and Comparative Law Quarterly 481, 483.

82 Case C-130/10 *Parliament v Council* (n 24).

83 Eeckhout suggests that if there was such a material difference between the two, then 'CFSP decisions ought not to be as detailed as they usually are': Eeckhout (n 43) 483.

the decisions taken under the CFSP since the Treaty are related to imposition of restrictive measures on third states or third parties. The principal exceptions are the decisions on humanitarian assistance in Libya[84] and reform of the security sector in the Democratic Republic of the Congo.[85] Therefore, the impression given by the Treaty that legislative acts are excluded since they are not needed or applicable to the domain of high politics is not borne out in practice.

Despite the lack of use of the CFSP instruments at their disposal, calls by Member States for the CFSP to take on a fully legislative character are few and far between. The report by the 'Group of Europe' of 11 Member State foreign ministers in their future-oriented plan for Europe in September 2012[86] listed as a priority, after resolving the euro crisis, strengthening the EU's act on the world stage. However, it did not call specifically for legislation in CFSP but rather majority voting in *decisions*. The report also noted the need to reduce the ability of a single Member State to block decisions but also to 'develop the concept of constructive abstention'.[87] What this tells us is that even the Member States who most favour integration as a means to make the EU work better stop short of including the CFSP wholeheartedly within the integration process. That is not necessarily to say that these Member States are immune to the prospect of greater co-operation, or even supranationalism, but that the institutional focus remains on the Council, where it is not just the Member States but the executives of the Member States which hold the reins.

The quotation at the outset suggests that the CFSP is subservient to the will of the Member States acting in a collective manner and completely separately from the other areas of EU integration where the supranational institutions are allowed to act. But if this was the case, then the CFSP would not be an 'equal area of action' but should take priority over all other externally focused areas. The practice shows that this is not the case. Rather, the CFSP serves a function for action on matters which fall within the CFSP as a residual category. If this were not the case, then there would be little achieved by inserting provisions which apply across the EU's external actions nor insisting on the need for coherence and consistency in external relations. Even then, the perspective of the CFSP as being intergovernmental is not only out-dated but misleading because it stresses that the Member States are the only significant actors in it and that anything which concerns the world beyond the borders of the EU must take place within CFSP. The critique of the ring-fencing does not, as would seem at first glance, demonstrate that a claim (official or otherwise) that the CFSP is more intergovernmental since Lisbon is justifiable in practice.

That the EU needs to act more coherently in external relations is almost universally agreed, even if the extent to which the EU should have a global role (and with what powers) is not. The 'face' of a more coherent CFSP is, according to the Treaty, the High Representative for CFSP; a role which was downgraded from the proposal in the Constitutional Treaty for a Union Minister for Foreign Affairs. Though modelled on an executive concept of foreign policy, the title was changed as part of the jettisoning of

84 Council Decision 2011/210/CFSP of 1 April 2011 on a European Union military operation in support of humanitarian assistance operations in response to the crisis situation in Libya (EUFOR Libya) [2011] OJ L89/17.

85 Council Decision 2010/576/CFSP of 23 September 2010 on the European Union police mission undertaken in the framework of reform of the security sector (SSR) and its interface with the system of justice in the Democratic Republic of the Congo (EUPOL RD Congo) [2010] OJ L254/33.

86 *Final Report of the Future of Europe Group of the Foreign Ministers of Austria, Belgium, Denmark, France, Italy, Germany, Luxembourg, the Netherlands, Poland, Portugal and Spain* (17 September 2012) <www.auswaertiges-amt.de/cae/servlet/contentblob/626322/publicationFile/171798/120918-Abschlussbericht-Zukunftsgruppe.pdf>.

87 Ibid 7.

constitutional/state-like features of the Constitutional Treaty. Nevertheless, the development of the pre-Lisbon post of High Representative who functioned as Secretary General of the Council merits consideration, since the new institutional arrangements do not bear out a CFSP which is effectively ring-fenced from other policy areas.

There are two reasons for this. The first is that the High Representative, inheriting her role from the pre-Lisbon High Representative, who was only a servant of the Council, is also a Commission Vice-President. This innovation is indicative of the porous nature of the institutional arrangements which is designed to promote coherence of action, but does not reflect the decision-making structure of the CFSP as laid down in the Treaties.[88] The Commission is therefore to be more involved in the CFSP than the 'specific provisions' would suggest. This is a strong element of the requirement of consistency, but also a paradox when it comes to ring-fencing: the High Representative is presumed to keep separate the respective roles and to be vigilant when matters discussed in the Commission pertaining to non-CFSP areas do not impinge on competences which should be within the CFSP. The 'double-hatted' High Representative is somehow expected to wear different intergovernmental/supranational hats depending on the circumstances. The post of High Representative does not constitute an EU institution in itself, although Article 18 TEU lists the post besides the other institutions, and thus the Treaty leaves room open to institutional practice to define where the fence lies.[89] The effectiveness of the CFSP ring-fencing is brought into question by the High Representative herself, who in her own words states that: 'my first impressions for the 18 months are that this [role of High Representative] is a huge role, created without deputies and created on paper without any reference to look back on of a description of how it would actually be in practice',[90] or indeed 'like flying a plane while you are still building the wings and somebody might be trying to take the tail off at the same time'.[91] This is partly as a result of the way in which CFSP came into the reform process, which was not because of difficulties in the way in which the High Representative's role worked but to enable institutional oversight over the drive for coherence which emerged during the reform process.[92] As such, neither the theory nor practice of this institutional innovation supports the ring-fencing of the CFSP.

The second is the development of the EEAS. Article 27(3) TEU defines the role of the EEAS to assist the High Representative and the principle of staffing it by officials drawn from the Council and Commission, and seconded from Member States.[93] This was a dramatic departure from the previous practice of external representation of the EU in third countries through Commission delegations and co-operation between EU/Member State officials without being physically situated in the same building. The Council Decision

88 I am grateful to Peter Van Elsuwege for helpful discussions on this point.

89 D Thym, 'The Intergovernmental Constitution of the EU's Foreign, Security and Defence Executive' (2011) 7 European Constitutional Law Review 453, 457.

90 House of Lords Select Committee on the European Union, Evidence Session with Baroness Ashton of Upholland, High Representative for Foreign Affairs and Security Policy, Vice-President of the European Commission, Evidence Session No 1, 14 June 2011, 3.

91 Ibid 4.

92 Sari (n 12) 79.

93 'In fulfilling his mandate, the High Representative shall be assisted by a European External Action Service. This service shall work in cooperation with the diplomatic services of the Member States and shall comprise officials from relevant departments of the General Secretariat of the Council and of the Commission as well as staff seconded from national diplomatic services of the Member States. The organisation and functioning of the European External Action Service shall be established by a decision of the Council. The Council shall act on a proposal from the High Representative after consulting the European Parliament and after obtaining the consent of the Commission.'

establishing the EEAS[94] requires all staff (which is to include a 'meaningful' presence of nationals from all Member States)[95] to act *only* in the interests of the EU.[96] The role of the EEAS is to assist in the fulfilment of the mandate relating to CFSP and other external competences belonging to the Commission and Council[97] and to help the High Representative 'ensure overall political coordination of the Union's external action, ensuring the unity, consistency and effectiveness of the Union's external action'.[98]

Although the Council Decision establishing the EEAS was based on a CFSP provision (Article 27(3) TEU), it appears from the procedural aspects of the way it was adopted that its nature and scope are not limited to the CFSP only.[99] In particular, the EEAS is not the servant of the Council but an 'inter-institutional service' and subject to influence exerted on it by the Commission too.[100] Here again, we see a holistic view of EU external relations being carried out in a way which should not fall foul of the EU's (internal) institutional divisions. But the institutional arrangements do not support the carrying out of EU external relations according to a strict separation of CFSP from non-CFSP matters. Despite the delicate 'balancing act' reflected in the Council Decision which founded the EEAS between the institutional conduct of CFSP versus other external relations,[101] in reality there are few conceptual or practical lines which can be drawn between what is (or should be) CFSP or not. In particular, whilst approximately one-third of the staff of the EEAS should be drawn from the diplomatic services of all the Member States,[102] the Council Decision does not imply that the work of the staff should be divided across CFSP and non-CFSP lines. Indeed, this would practically be impossible in most cases given the wide definition of CFSP aims, but policies that exist under other competences.

Whether the influence of the intergovernmental CFSP within the conduct of external relations by the High Representative permeates other policies, or vice versa, will be evident over time and will depend on a variety of factors (including the working culture of the section of the EEAS, the delegation in a third country or the nature of the relationship with the third country or issue under question). Since all EEAS staff are required to work in the interests of the EU, this suggests a sidelining of the CFSP given the extensive external assistance instruments in place in relation to many third countries.[103] Previous work on the socialisation of officials working in European institutions, including pre-Lisbon CFSP, suggests that CFSP is already much less 'intergovernmental than this term allows'[104] and,

94 Council Decision 2010/427/EU of 26 July 2010 establishing the organisation and functioning of the EEAS.

95 Ibid preamble point 8.

96 Ibid preamble point 9.

97 Ibid Article 2.

98 Ibid Article 9(2).

99 S Blockmans and C Hillion (eds), *EEAS 2.0: A Legal Commentary on Council Decision 2010/427/EU Establishing the Organization and Functioning of the European External Action Service* (Centre for European Policy Studies 2013) 2.

100 M Furness, 'Who Controls the European External Action Service? Agent Autonomy in EU External Policy' (2013) 18 European Foreign Affairs Review 103, 112.

101 J Wouters, G De Baere, B Van Vooren, K Raube, J Odermatt, T Ramopoulos, T van den Sanden and Y Tanghe, 'The Organisation and Functioning of the European External Action Service: Achievements, Challenges and Opportunities' (2013) European Parliament, Directorate-General for External Policies Study EXPO/B/AFET/2012/07, 25.

102 Council Decision 2010/427/EU, Article 6(9).

103 Ibid Article 9.

104 A Juncos and K Pomorska, 'Invisible and Unaccountable? National Representatives and Council Officials in EU Foreign Policy' (2011) 18 Journal of European Public Policy 1096; and A Héritier, *Explaining Institutional Change in Europe* (Oxford University Press 2007) 6–7.

given the longevity of co-operation already existing within the institutions on foreign policy, has already evolved into something akin to law.[105]

The role of the Commission goes beyond the High Representative herself. It would be futile to suggest that, even before Lisbon, the Commission was wholly divorced from the workings of the CFSP. Indeed, as former External Relations Commissioner Chris Patten has himself said, a great deal of his time was devoted to the CFSP.[106] If the ring-fencing in the Treaty was supposed to limit the supranational influence of the Commission as well as the Court of Justice and the Parliament, then this is a message which has not reached the Commission itself. Indeed, since the entry into force of the Treaty of Lisbon, a new inter-institutional agreement between the European Parliament and the Commission foresees the involvement of the former by the latter in the CFSP: 'Within its competences, the Commission shall take measures to better involve Parliament in such a way as to take Parliament's views into account as far as possible in the area of the Common Foreign and Security Policy.'[107]

If the Commission was intended to play no role in the post-Lisbon CFSP, then this provision would be meaningless. Rather, it is demonstrative of the Commission's informal role in the CFSP and the inadequacy of the Treaty provisions to account for or reflect what actually occurs in the CFSP, or external relations more generally. A similar point has been raised in relation to the European Parliament, which has more indirect influence over the institutional workings of the CFSP than is commonly assumed.[108] These observations are closely linked to the need for coherence, which in itself implies unitary action, despite the legal 'fences' put in place. It does not seem likely that the Court is seen as a more supranational 'threat' than the Commission. Rather, it can be said that the Commission is uniquely able to promote the coherence of EU external relations since it is responsible for the non-CFSP common commercial policy. At the one level, CFSP should concern 'political' foreign policy issues which are not, or cannot be, dealt with via binding legal measures. But this distinction had been shown to be unworkable before Lisbon. Post-Lisbon, with the requirement to ensure consistency, is the CFSP to be reduced to 'purely' political issues which are ring-fenced? As stated above, under the previous arrangements, if there was a doubt over which legal basis a measure should be based upon, then the general answer was clear.[109] But add to this the increasing number of examples of where 'cross-pillar' policies took place, such as the European Neighbourhood Policy (ENP), which have transcended the formal divisions in the Treaties and which the High Representative (for CFSP) emphasises as a priority area. Given the requirement for consistency in the Treaty, it seems likely that, should the EU continue to seek for coherent policies towards third countries/regions/issues, wide-ranging policies which are not simply based on one particular legal basis will become the norm.

This point is supported by an analysis of the decisions which have been taken under the CFSP since the entry into force of the Treaty of Lisbon. Despite the all-encompassing foreign affairs the Treaty is supposed to cover, in fact the vast majority of outputs from the CFSP are concerned with economic sanctions on third states and individuals. The *Kadi*

105 M E Smith, *Europe's Foreign and Security Policy: The Institutionalization of Cooperation* (Cambridge University Press 2004) 117.

106 C Patten, *Not Quite the Diplomat* (Penguin 2005) 16.

107 Framework Agreement on Relations between the European Parliament and the European Commission (2010) 20.11.2010 OJ 304/47, Article 10.

108 E Wisniewski, 'The Influence of the European Parliament on the European External Action Service' (2013) 18 European Foreign Affairs Review 81.

109 Brkan (n 36) 101.

situation would still arise, since it affects legal rights of third parties. Furthermore, Article 215 TFEU puts in place measures agreed under 'general' policy of CFSP.[110] The two logical conclusions to draw from this observation are that (again, given the wide-ranging nature of the CFSP provisions) EU foreign policy is redundant, or that it is alive and well, but the instruments under CFSP are not being used. The contention here is that it is the latter.

Ring-fencing risks effectively restricting the CFSP's scope to a 'nucleus' of purely 'political' measures. If so, there are two potential consequences. First, that the CFSP could be used for a general agreement on a (possibly) high-profile international issue on which the Member States are agreed and which to announce their common view to the world. This would provide a springboard to further actions taken under other EU competences, and would reflect the line of reasoning from *Parliament v Council*.[111] Second, the CFSP would, with the exception of the sanctions measures, continue to become even more declaratory in nature, meaning that decisions taken under CFSP would simply give an indication as to the level of agreement existing between the Member States on a certain issue. But the problem with this assertion is that it suggests that a neat distinction exists between 'high politics' and 'low politics', which is superficial.[112] If it was the case that the Member States had pooled aspects of foreign policy such as recognition of new states, then the CFSP could be used in this way. But as recent examples show, the declaration of independence of South Sudan, or the recognition of the Libyan or Syrian rebel forces as the legitimate governments of those countries, Member States have not taken the opportunity collectively. An alternative, which recalls Lavenex's 'concentric circles' of EU external governance, is that the CFSP is retained for measures on the very outer limit of what Member States are prepared to commit themselves.[113] *Parliament v Council*[114] can be read as supporting the view that CFSP is simply there to 'scope' general aims and then flesh out the actions via other competences.

The Council also maintains that it does not follow from the fact that listed persons and entities may now bring an action for the annulment of decisions taken in the sphere of the CFSP imposing restrictive measures on them that any amendment to an existing regulation must necessarily be preceded by the adoption of a new CFSP decision.[115]

And yet, the argument of the CFSP being at the centre of foreign policy from which other initiatives flow does not support this proposition. According to the High Representative, the foreign policy 'core' is provided by the ENP and 'recognising the importance of the bilateral connections that we have' in strategic partnerships with the USA, Russia, China, India, Brazil or South Africa.[116] In evidence given to the House of Lords EU Select Committee, the High Representative does not even mention the CFSP.[117] Rather, the 'priority' relationships are not defined by CFSP instruments and the lack of use of common strategies pre- and post-Lisbon suggests that Member States do not foresee the use of CFSP in this way. Therefore, if these core relationships and their legal bases are found elsewhere, then the ring-fencing does not suggest that the CFSP occupies a 'higher' place in the legal hierarchy from which initiatives may flow. There has been no call for a

110 Eckes (n 74) 120.

111 Case C-130/10 *Parliament v Council* (n 24).

112 Van Elsuwege (n 59) 988.

113 S Lavenex, 'Concentric Circles of Flexible "EUropean" integration: A Typology of EU External Governance Relations' (2011) 9 Comparative European Politics 372.

114 Case C-130/10 *Parliament v Council* (n 24).

115 Ibid para 99.

116 House of Lords Select Committee (n 90) 8.

117 Ibid.

re-evaluation of existing external policies on the grounds of their legal basis. If Member States are responsible for ring-fencing the CFSP in this way, and assuming that the UK is the 'sceptical' benchmark for any integration of foreign policy, then the UK's enthusiasm for EU-led external initiatives such as an EU–US free-trade agreement[118] means that a link between the ring-fenced CFSP and all other areas is not what was intended. Rather, the non-CFSP competences will continue to develop of their own accord whilst the CFSP is left as a residual character: something which is not entirely reflected in the words of the Treaty.

The reasoning set out above leads to a conclusion that the practice of the CFSP, beyond sanctions, remains declaratory in nature. 'Declaratory' is a criticism that has been levelled at the CFSP since its creation, and whilst declarations may have some foreign policy impact, it is curious that these are the hallmark of the policy, instead of the instruments which have been specifically created for its use. The extent to which non-CFSP measures are used already suggests that actions and policies toward third countries or issues are there but not badged as such under the CFSP. The ENP is good example of this, as a document issued jointly by the Commission and High Representative notes that CFSP engagement 'will continue to be part and parcel of the ENP'[119] and the roadmaps of actions to take within the framework of the ENP do not mention the use of CFSP instruments.[120] This seems to be at odds with Article 29 TEU which suggests that the Union's approach to a particular geographical or thematic issue will be taken via a CFSP decision and Article 40 TEU. A better way to explain the ring-fencing with reference to the Member States is, rather than to suggest that they are immune to foreign policy co-operation (or even/rather integration), that the CFSP can remain declaratory but free of 'legalism'. This is also suggestive of a continued development of EU external relations which does not depend on Treaty changes as a catalyst but rather incremental, institutional development.[121] Taken together, these observations place the CFSP in a category of executive actions for which a court (here, the Court of Justice) would be unable to find any binding characteristics. The immunity of foreign policy from judicial control is reflected in the national constitutional arrangements of Member States. A strong argument against democratic oversight would be that there is no (legal) substance to a purely political policy area and the previous sections in this article have demonstrated that strictly separating what is or should be CFSP or not is, at best, extremely difficult and liable to focus attention on internal divisions with the EU rather than the co-ordinated common foreign policy the CFSP is designed to further.

Another consequence of a declaratory CFSP is that, in terms of both legal certainty and meeting the aims of the Treaty, the use of the CFSP provisions itself is only going to have a marginal effect on the development of EU foreign policy. This is also regrettable because the two 'new' features within the Treaty relating to foreign policy (coherence and the promotion of values) suggest that the EU is going to be measured according to these as well as its general capacity to forge a common policy. The focus of attention is likely to remain on what is being done under the CFSP rather than anything else. By attempting to ring-fence the CFSP in this way, in a way which appears theoretically possible but practically

118 David Cameron, speech to the World Economic Forum in Davos, Switzerland, 24 January 2013 <www.number10.gov.uk/news/prime-minister-david-camerons-speech-to-the-world-economic-forum-in-davos/>.

119 Commission/High Representative of the European Union for Foreign and Security Policy, *European Neighbourhood Policy: Working Towards a Stronger Partnership* JOIN (2013) SWD (2013) 79 final.

120 Commission/High Representative of the European Union for Foreign and Security Policy, *Eastern Partnership Roadmap 2012–2013* JOIN(2013) SWD (2013) 108 final.

121 M Smith, 'Still Rooted in Maastricht: EU External Relations as a Third-Generation Hybrid' (2012) 34 European Integration 699, 704.

unworkable, the CFSP itself risks being marginalised to the extent that even the High Representative, when characterising the priorities of European foreign policy, does not seem to identify it as a means by which EU foreign policy can be led, but rather the opposite.[122]

6 Conclusion

Major EU treaty negotiations have become ever more cumbersome, complex and time-consuming as the political demands of Member States translate into legal provisions which make the EU constitutional animal an even stranger beast. No more so is this the case than for the CFSP. The reforms brought about by the Treaty of Lisbon do two conflicting things: strengthening the visibility and capabilities of the EU to act on the international plane in a more consistent and coherent manner, whilst also at least purporting to set clearer institutional lines which cannot be crossed, but which, if they could be crossed, would be an integral part of helping the EU achieve its goals based on the EU's experience in other areas. If the provisions of the CFSP can be characterised as more 'political' than 'legal', then this suggests that the intention of the Treaty of Lisbon reforms is to ring-fence the CFSP around the most sensitive areas in terms of state sovereignty. This has a dual function; on the one hand, those Member States concerned with selling the Treaty to (sceptical) domestic populations could point to the ring-fencing as a form of protection of national sovereignty away from the integrationist EU institutions. On the other hand, it allows the competences of the rest of the Treaty provisions to be used for the actual conduct of the EU's external relations. This article has also shown that, rather than the CFSP being the starting point for EU policy towards a thematic or geographical issue in external relations, in practice it is unlikely to be used as such for initiatives in foreign policy which can then be followed up by using other Treaty competences.

Ring-fencing to ensure that the Member States are the only significant actors on key areas from which all else follows is not borne out in the institutional practices permitted (or even promoted) under the Treaty. In the 'real' exercise of the EU's external action, a more complex picture emerges. From a strict rule of law perspective, one might consider that this frustrates the aims of the drafters of the Treaty since the letter of the law is not being obeyed if the CFSP instruments are ignored in favour of measures found elsewhere to pursue what should be CFSP goals. From the perspective of consistency and coherency of EU external action, this practice might be a welcome development in terms of satisfying the general goals of the Treaty. As a result, the boundaries between CFSP and other areas are likely to continue to provide a fertile ground for analysis by legal scholars, who, like the Court of Justice, must at the same time unpick the political reasoning behind the paradox of a ring-fenced CFSP, which is more porous than it would seem. To fully understand where the law in CFSP is, scholars must continue to look beyond the Treaty-based instruments and see where the goals are fulfilled via other means.

However, the implications for the nature and role(s) of law in relation to the EU more generally are not limited to arguments based on competences. Rather, the position of the CFSP in the EU's Treaty arrangements and attempt to ring-fence is indicative of an attempt to further remove this policy area from the influence of what has become 'normal' in EU law and what is known to be effective in making the law work across so many different Member States. The Treaty's reforms tell us that there is a fear of EU law applying to this area beyond almost all others and that the possibilities for further integration by law in this area may have passed. The reticence to allow, within the text of the Treaty, the supranational institutions to carry out the work they have done in areas of European integration seen as

122 House of Lords Select Committee (n 90).

most successful (such as the internal market) pushes the CFSP back towards the realm of classic international law. The problem of doing so, as is well known, is that decision-making processes and the decisive action on world affairs so frequently called for are unlikely to be forthcoming within this type of legal framework.

Since some Member States publicly support the greater use of majority voting in CFSP (though not necessarily the extension of the 'normal' legislative decision-making process), it might be thought possible that CFSP becomes part of a 'multi-speed' or differentiated pattern of integration where certain Member States integrate further. But insofar as the CFSP represents the view of the Union as an emerging international actor, it is surely in a different situation from a policy-making area which applies only internally within the EU. The overall effectiveness argument, whilst certainly present whenever multi-speed Europe is discussed, is more significant in the CFSP since its *raison d'être* is to enable the EU (and not only a part of it) to be a more effective *global* actor and therefore its significance relies on the unity and collective weight of all its Member States. The consequence for the EU's legal order is that it would be difficult to see how the aims of consistency and coherence could be met with differentiated patterns of integration in this area, and hence possibly other areas too.

Conversely, in the wider debates about what the EU should be, how far (and in what areas) it should integrate and what type of 'law' is fit for purpose, ring-fencing other areas might prove attractive to Member States seeking to 'repatriate' powers away from the supranational institutions. This might conceivably be in the areas of freedom, security and justice; migration (especially insofar as third-country nationals are concerned); or even the internal market. If this does occur, then it may be that the traditional method of legal integration via regulations and directives loses ground to a model of EU law where Member States ring-fence certain areas. The result would be a fragmentation of the EU's legal order and the calling into question of one of the most fundamental dimensions of the 'new legal order' which has been at the core of the integration process.

NILQ 64(4): 465–83

Twenty years on, the right of silence and legal advice: the spiralling costs of an unfair exchange

HANNAH QUIRK*

University of Manchester

Abstract

This article appraises the cumulative effects of the curtailment of the right of silence after 20 years. A controversial measure when enacted, its long-term effects have not been examined. It is argued that the Act was introduced on a flawed premise; has been interpreted very broadly; and has facilitated further restrictions on due process rights. Drawing upon the case law, academic commentary and empirical research, it is argued that the provisions have compromised the lawyer–client relationship and undermined the effectiveness of legal representation. Not only have the provisions made police interviews part of the trial (as Jackson argues), but they have also imported to the courtroom traditional police suspicions of defendants and their lawyers. The court has accordingly interpreted the provisions to the detriment of another 'fundamental condition'; legal professional privilege. The changes have thus had a wider impact than was envisaged at the time.

Key words: right of silence; police station legal representation; legal professional privilege; Court of Appeal

The right of silence and the legal representation of suspects had a special status in the criminal justice system, described respectively as one the 'sacred cows'[1] of British justice and 'one of the most important and fundamental rights of a citizen'.[2] Both rights were also viewed with suspicion by some as a means by which 'criminals' could frustrate justice. Throughout most of the twentieth century, with some caveats, suspects were entitled to withhold their accounts from the police and also to decline to testify at trial;[3] a right controversially curtailed by ss 34–38 of the Criminal Justice and Public Order Act 1994 (CJPOA). The approaching 20th anniversary of this change, with concerns again being voiced

* Lecturer in Criminal Law and Justice, University of Manchester. I would like to thank David Hamer, Leonard Leigh, Roger Leng, Kieran McEvoy, John Murphy and Paul Roberts for their encouragement, advice and comments on drafts of this article.

1 A R N Cross, 'The Right to Silence and the Presumption of Innocence: Sacred Cows or Safeguards of Liberty' (1970–1971) 11 Journal of the Society of Public Teachers of Law 66–75.

2 *R v Samuel* [1988] 87 Cr App R 232, 245.

3 See P Mirfield, *Silence, Confessions and Improperly Obtained Evidence* (Oxford University Press 1998).

about the possibility of miscarriages of justice, this time caused by the proposed cuts in legal aid,[4] seems an appropriate juncture to examine the long-term effects of these changes.

The CJPOA permits prosecutorial and judicial comment and the drawing of 'such inferences as appear proper' when a defendant:

- relies at court upon any fact not mentioned during questioning or charge, that could reasonably have been raised at that stage (s 34);
- fails to account for the presence of any substance, object or mark on or in the suspect's person, clothing, footwear or possession (s 36); or his or her presence at a particular place (s 37);
- fails to testify or, having been sworn, refuses to answer questions without good cause (s 35).

The right of silence is often used, as here, as a convenient, if incomplete shorthand. 'It is important to remember that the significance of section 34 does not lie in silence in interview, it lies in reliance at trial on something that should have been said in interview.'[5] A police officer is 'entitled to question any person from whom he thinks useful information can be obtained . . . A person's declaration that he is unwilling to reply does not alter this entitlement.'[6] A suspect can be detained to 'secure or preserve evidence relating to an offence for which he is under arrest or to obtain such evidence by questioning him'.[7] The police may take arrested suspects into custody in order to increase the likelihood of their answering questions,[8] but suspects did not,[9] and indeed do not,[10] have to answer police questions. Suspects are cautioned when arrested and at the start of every interview: 'You do not have to say anything. But it may harm your defence if you do not mention when questioned something which you later rely on in Court. Anything you do say may be given in evidence.'[11]

The changes made by the CJPOA were justified partly as a trade-off against the improved rights provided for suspects by the Police and Criminal Evidence Act 1984 (PACE), in particular, the entitlement to free and independent legal advice at the police station. Although the right to consult a solicitor at an early stage of the investigation existed at common law, very few suspects actually exercised this right before PACE. Greer characterised this trade-off between the right of silence and improved protections as 'Exchange Abolitionism',[12] but this article argues that such an exchange was not fair and has cost suspects far more than was envisaged at the time.

4 'Legal Profession Joins Forces to Oppose Unreasonable Legal Aid Proposals' press release 14 May 2013 <www.barcouncil.org.uk/media-centre/news-and-press-releases/2013/may/legal-profession-joins-forces-to-oppose-unreasonable-legal-aid-proposals/> accessed 25 May 2013; D Pannick, 'Why Chris Grayling Should Listen to the Evidence on Legal Aid: Miscarriages of Justice then Create Further Costs to the State' *The Times* (20 June 2013).

5 *R v Essa* [2009] EWCA Crim 43, [15].

6 PACE Code C, note 1K.

7 S 37(2) PACE.

8 *Holgate-Mohammed v Duke* [1984] AC 437; PACE Code C, Notes for Guidance 1B.

9 *Rice v Connelly* [1966] 2 QB 414.

10 There is no requirement to answer questions as there is, for instance, under pt XIV of the Companies Act 1985.

11 PACE Code C, 10.5.

12 S Greer, 'The Right to Silence: A Review of the Current Debate' (1990) 53(6) MLR 709–30.

For such a contentious measure, there has been little investigation of the effects of the curtailment of the right of silence,[13] and there appears little appetite for its restoration.[14] Academic analysis focusing on the case-law initially described the changes as being of only 'marginal' evidential significance,[15] or of greater burden than benefit to the prosecution.[16] Little attention has been paid to the topic since the case-law became more settled. This contrasts with the impassioned debate when the CJPOA was enacted, when the right of silence was attacked for allowing experienced criminals to frustrate justice, and defended both as a protection against miscarriages of justice and as a signifier of the presumption of innocence. As Leng suggests,[17] the almost exclusive focus on these polarised positions meant that insufficient attention was paid to other theoretical and practical problems which might result from the CJPOA. In particular, inadequate consideration was given to the effect that allowing inferences to be drawn from no-comment interviews at the police station would have on the protective value of custodial legal advice, and the relationship between suspects and their legal representatives.[18]

This article argues that suspects have been trebly disadvantaged by the CJPOA. The right of silence has been diminished, legal representation has been devalued, and the encroachment on such 'a "bench mark" of British justice'[19] facilitated further statutory and common law restrictions on evidential protections for defendants. The argument is primarily based on the case-law and literature but it also draws illustrative quotes from the findings of a qualitative study conducted in the early days of the CJPOA.[20] Part one explains how the CJPOA was based upon the false premise that prosecutions and convictions were being thwarted by legal advisers advising their guilty clients to make no-comment interviews. Such criticisms were expressed at the time, but are summarised here in order to illustrate the full trajectory of the changes and the unfair foundations of the process. Part two examines how a largely unsympathetic Court of Appeal has developed the scope of the provisions in accordance with the exchange-abolitionist viewpoint and by ignoring the practical difficulties faced by legal representatives at the police station. It is argued that this has compromised the lawyer–client relationship – beyond Cape's suggestion of defence lawyers being merely 'sidelined'[21] – and undermined legal representatives' ability

13 The only published empirical study to date is T Bucke, R Street and D Brown, *The Right of Silence: The Impact of the Criminal Justice and Public Order Act 1994* (Home Office Research Study 1999, Home Office 2000).

14 Cl 11 of the Liberal Democrats' draft Freedom Bill 2009 which proposed the 'repeal of provisions which restrict the right to silence' attracted little attention (the document has been removed from the Liberal Democrats' website; a copy is on file with the author). The government decided not to include it in the Protection of Freedoms Bill 2010–2011 because the current arrangements 'strike the right balance' (email correspondence between the author and the Ministry of Justice, dated 1 August 2011).

15 I Dennis, 'Silence in the Police Station: the Marginalisation of Section 34' [2002] Crim LR 25–38.

16 D Birch, 'Suffering in Silence: A Cost-Benefit Analysis of Section 34 of the Criminal Justice and Public Order Act 1994' [1999] Crim LR 769–88.

17 R Leng, 'The Right to Silence Reformed: A Re-appraisal of the Royal Commission's Influence' (2001) 6(2) Journal of Civil Liberties 107–33, 125.

18 Legal representation at the police station must be provided by either a solicitor who holds a current practising certificate, or an accredited or probationary representative included on the register of representatives maintained by the Legal Services Commission (para 6.12 Code C PACE). The terms 'solicitor' 'lawyer' and 'legal representative' are used interchangeably in this article but not all those acting at the police station are fully qualified.

19 J D Jackson, 'Curtailing the Right of Silence: Lessons from Northern Ireland' [1991] Crim LR 404–15, 415.

20 100 semi-structured interviews were conducted with criminal justice practitioners from five differently sized locations in one large metropolitan region of England between 1998 and 2001 (legal representatives, barristers, crown prosecutors, magistrates, legal advisers and judges). Each lasted around one hour and was fully transcribed and coded. An additional 100 questionnaires were completed by police officers (out of 300 issued).

21 E Cape, 'Sidelining Defence Lawyers: Police Station Advice after *Condron*' (1997) 1 E&P 386–402.

to protect their clients at the police station. The CJPOA has created an expectation that the accused will co-operate with proceedings against them, in many cases making it easier for the prosecution to discharge the burden of proof. Jackson's argument that the CJPOA provisions have effectively made the police interview part of the trial process[22] is expanded upon to suggest that the provisions have imported to the courtroom traditional police suspicion of defendants and their legal representatives. Part three of the article explores how, in accordance with this viewpoint, the courts have interpreted the CJPOA provisions to the detriment of another 'fundamental condition', that of legal professional privilege. Despite the apparent acceptance of the status quo, it is this 'ripple effect' of the legislation that calls for further examination of the issue. Leng argues that the focus on implementing the law 'must not obscure the extent to which the current law compromises the core notions of adversarial justice and fair trial'.[23] When considered holistically, it is argued that the cumulative effects of the legislation have been corrosive to many aspects of the adversarial process but, in particular, and the focus of this article, on the right to legal representation at the police station.

1 Exchange abolitionism: part one of a three-card trick

Opposition to the right of silence is not new; Jeremy Bentham's criticism of it as the first law a team of criminals would draft are cited routinely by abolitionists.[24] Legislation to curtail the right of silence was drafted by the Criminal Law Revision Committee in 1972[25] but was shelved following formidable opposition.[26] The next major review of the system by the 1981 Royal Commission on Criminal Procedure carefully balanced increased powers for the police against improved protections for suspects and explicitly included the right of silence as part of that equation.[27] Its recommendations were largely given effect by PACE which provides miscellaneous requirements to ensure minimum standards of detention, that interviews are conducted properly and are tape-recorded. Of particular significance for this debate, suspects must be informed that they have the right to free, independent legal advice, in private, by telephone, in writing or in person, which may be delayed only in exceptional, limited circumstances.[28] The function of the legal representative is solely 'to protect and advance the legal rights of his client'; this includes explaining, and sometimes advising, the exercise of the right of silence.[29]

The post-PACE regime was portrayed by some politicians as assisting 'professional criminals, hardened criminals and terrorists who disproportionately take advantage of and

22 J Jackson, 'Silence Legislation in Northern Ireland: The Impact after Ten Years' (2001) 6(2) Journal of Civil Liberties 134–49.

23 Leng (n 17) 133.

24 Cross (n 1) 73; Criminal Law Revision Committee, *Eleventh Report: Evidence (General)* (Cmnd 4991, 1972) 18; G Williams, 'The Tactic of Silence' (1987) 137 NLJ 1107; P Imbert, 'Policing London' (1988) 61 Police Journal 199, 203; I McKenzie and B Irving, 'The Right of Silence' (1988) 4 Policing 88, 91–92; Michael Bates MP HC Deb 7 July 1993, vol 228, col 334.

25 Criminal Law Revision Committee (n 24).

26 M Zander, 'The CLRC Report: A Survey of Reactions' (1974) LSG (7 October); P Evans, 'Criminal Law Revision Proposals Received with Heavy Criticism' *The Times* (28 June 1972) 1.

27 Royal Commission on Criminal Procedure, *Report* (Cmnd 8092, 1981).

28 See Code of Practice C, Annex B. A less favourable regime is in place for suspects detained under s 41 Terrorism Act 2000; this is governed by Code H.

29 PACE Code C, 6D.

abuse the present system',[30] and the police 'waged a concerted campaign'[31] to this effect. Following PACE, the Court of Appeal was generally supportive of legal representation ('in marked contrast to the pre-1984 attitude'),[32] but appeared to regard this as coming at a price. First, it took the view that the presence of a legal representative put suspects on 'equal' or 'even' terms with the police,[33] meaning that improperly obtained evidence need not always be excluded. It then went further in aiming 'to secure a quid pro quo for the enforcement of section 58 rights [to legal advice], in terms of a new right to comment adversely on a suspect's failure to reveal a defence later sprung on the prosecution'.[34] Some senior judges and prosecutors also expressed this view in a personal capacity.[35]

In 1988, a Home Office Working Group[36] was established to report how, rather than whether or not, the right of silence should be curtailed. The right was 'removed with unseemly despatch'[37] in Northern Ireland, ostensibly following a number of terrorist incidents.[38] Following a series of terrorism-related miscarriages of justice, reforms were deferred in England and Wales until the Royal Commission on Criminal Justice (RCCJ)[39] was controversially asked to consider the question once again. The majority of the RCCJ concluded that:

> the possibility of an increase in the convictions of the guilty is outweighed by the risk that the extra pressure on subjects to talk in the police station and the adverse inferences invited if they do not may result in more convictions of the innocent.[40]

The government rejected the advice of the RCCJ and introduced the CJPOA, beginning an era of criminal justice policy based on the premise that: 'The balance in the criminal justice system is tilted too far in favour of the criminal and against protecting the public.'[41] Such a policy appeared politically driven[42] rather than evidentially based. It is thus perhaps unsurprising that the legislation had unforeseen consequences. Dennis found that 'none of the claims which have driven the reform of the right to silence have any very secure

30 Michael Howard, HC Deb 11 January 1994, vol 235, col 26. The Secretary of State for Northern Ireland had commented similarly when curtailing the right of silence in Northern Ireland a few years earlier (Tom King, HC Deb 8 November 1988, vol 140, col 184).

31 J Jackson, 'Silence and Proof: Extending the Boundaries of Criminal Proceedings in the United Kingdom' (2001) 5 E&P 145–73, 145.

32 A A S Zuckerman, *The Principles of Criminal Evidence* (Clarendon Press 1989) 340.

33 *R v Chandler* [1976] 3 All ER 105, 110; *R v Smith* (1985) 81 Cr App R 286, 292.

34 D Birch, 'Case Comment, Confession: Denial of Access to a Solicitor' [1988] Crim LR 449–52, 452. See also Zuckerman (n 32) 327 and cases such as *R v Alladice* (1988) 87 Cr App R 380, 385; *R v Gilbert* (1978) 66 Cr App R 237.

35 Lord Denning, 'Free Justice from Silence' *The Sunday Times* (20 September 1987); Lord Chief Justice Taylor, HL Deb 23 May 1994, vol 555, col 519; and the 'Tom Sergant Memorial Lecture' (1994) 144 NLJ 125.

36 HC Deb 18 May 1988, vol 133, cols 465–66W; Home Office, *Report of the Working Group on the Right to Silence* (HMSO 1989).

37 Greer (n 12) 710.

38 Criminal Evidence (Northern Ireland) Order 1988 SI 1988/1987. See J Jackson, 'Recent Developments in Criminal Evidence' (1989) 40 NILQ 105–30.

39 RCCJ, *Report* (Cm 2263, 1993). See point (v) of its terms of reference.

40 Ibid 54.

41 Michael Howard, HC Deb 11 January 1994, vol 235, col 25.

42 Birch suggests that 'it is likely that section 34 owes its existence not so much to a considered choice between opposing views, as to the political incentive to reflect a change already made in Northern Ireland' (n 16) 772 (footnote omitted). It was also part of a change in policy set out in Michael Howard's 27-point plan announced to the Conservative Party Conference in October 1993.

foundation'.[43] Empirical evidence was available at the time that showed the right of silence was not exercised widely and was not insurmountable in achieving convictions.[44] Brown[45] reviewed the research to produce a 'best estimate' of the extent of the exercise of the right. He found that 5 per cent of suspects remained completely silent to police questions and between 6 per cent and 10 per cent remained partially silent. Leng's report for the RCCJ found no evidence that the right of silence was impeding prosecutions, or that no-comment interviews led to ambush defences at court.[46]

The exchange-abolitionist view of legal representation also favoured instinct over evidence. The right to legal advice at the police station has been described as an essential feature of a fair trial.[47] Article 6(3)(c) of the European Convention on Human Rights (ECHR) states that everyone charged with a criminal offence has the right 'to defend himself in person or through legal assistance of his own choosing or, if he has not sufficient means to pay for legal assistance, to be given it free when the interests of justice so require'. Research available at the time showed that less than one-third of suspects[48] was legally represented at the police station, that there was a 'lack of expertise and experience'[49] amongst many representatives, and that legal advisers were rarely as combative as they were claimed to be by the sponsors of the Act.[50] Many legal advisers regarded themselves as neutral referees or 'babysitters'[51] and did not want to risk prejudicing their working relationship with the police by advising no-comment interviews.[52] As one said:

> You've got to be lenient with the police, if you start rocking the boat with the police, then when you want some co-operation, you might not get it and at the end of the day, we're all there for the same thing aren't we – one side or the other side?[53]

The protections offered by PACE were undoubtedly an improvement on the common law, and the accreditation scheme for police station legal advisors has improved standards.[54] Suspects cannot be on 'even terms' with the police, however, when officers do not have to make full disclosure of their case, nor can a solicitor ameliorate the psychological effects of

43 I Dennis, 'The Criminal Justice and Public Order Act 1994: The Evidence Provisions' [1995] Crim LR 4–18, 14.

44 R Leng, *The Right to Silence in Police Interrogation: A Study of Some of the Issues Underlying the Debate* (Royal Commission on Criminal Justice Research Study No 10, HMSO 1993); M McConville and J Hodgson, *Custodial Legal Advice and the Right to Silence* (HMSO 1993).

45 D Brown, 'The Incidence of Right of Silence in Police Interviews: The Research Evidence Reviewed' in *Research Bulletin* No 35 (Home Office Research and Statistics Department 1994) 57–75.

46 Leng (n 44).

47 *Murray (John) v UK* (1996) 22 EHRR 29; *Salduz v Turkey* (2008) 49 EHRR 421; *Cadder (Peter) v HM Advocate* [2010] UKSC 43.

48 D Brown, T Ellis and K Larcombe, *Changing the Code: Police Detention under the Revised PACE Codes of Practice* (Home Office Research Study No 129, Home Office, 1992).

49 J Hodgson, 'Tipping the Scales of Justice: The Suspect's Right to Legal Advice' [1992] Crim LR 854–62, 861; D Dixon, A K Bottomley, C A Coleman, M Gill and D Wall, 'Safeguarding the Rights of Suspects in Police Custody' (1990) 1 Policing and Society 115–40; M McConville, J Hodgson, L Bridges and A Pavlovic, *Standing Accused: The Organisation and Practices of Criminal Defence Lawyers in Britain* (Oxford University Press 1994) 41; D Brown, *PACE Ten Years On: A Review of the Research* (Home Office Research Study No 155, HMSO1997).

50 J Baldwin, *The Role of Legal Representatives at the Police Station* (RCCJ Research Study No 3, HMSO 1993); McConville et al (n 49).

51 Solicitor/B2/i; J Hodgson, 'Adding Injury to Injustice: The Suspect at the Police Station' (1994) 21 Journal of Law and Society 85–101, 94; J Baldwin, 'Legal Advice in the Police Station' (1992) NLJ 1762–64, 1764.

52 McConville and Hodgson (n 44) 89.

53 Legal Executive (P)/F7/i); see also D Dixon, 'Common Sense, Legal Advice and the Right of Silence' (1991) PL 233–54, 246.

54 L Bridges and S Choongh, *Improving Police Station Legal Advice* (Law Society and Legal Aid Board 1998).

detention.[55] Although the police seem more comfortable with the presence of legal advisers nowadays, interviews take place on their territory and, to an extent, on their terms.[56] Research published before the CJPOA was enacted showed that the police were using 'tactics, or ploys'[57] to mitigate or minimise the effects of PACE. Officers have a wide discretion in the exercise of their powers, whereas legal representatives have little authority to enforce police compliance with the rules. The refusal to answer questions was hitherto the main bargaining chip that suspects and legal advisers possessed in response to the police exploiting, exceeding or abusing their powers.[58]

The CJPOA was based on the extraordinary syllogism that, if legal advisers were recommending that suspects exercise their rights, or were inhibiting the police from infringing these rights, then one of the most symbolic protections for suspects should be curtailed in order to 'rebalance' the system. There was no evidential basis for these claims, but much to contradict them. Whilst the Act was ostensibly targeted at the undeserving who were 'exploiting' the protections introduced to protect the vulnerable, its provisions apply to all, whether or not they have legal representation, and it is noteworthy that still fewer than half of suspects do.[59]

This unfair exchange has not settled the 'balance' between suspect and the police; the CJPOA appears rather to have marked the legislative start of 'the modern law and order arms race'.[60] Leng has argued that the CJPOA 'represents a re-negotiation of the relationship between citizen and state'.[61] Cape considers that '[a]ttempts to circumscribe the newly invigorated role of defence lawyers' were not limited to the CJPOA provisions but 'whether as a result of design or consequence'[62] may be discerned in other Acts. Further legislation has been passed encroaching on due process protections to an extent that would previously have been thought impossible, such as the creation of pre-trial disclosure obligations for the defence; exceptions to the double jeopardy rule; and the admission of bad character and hearsay evidence.[63] The CJPOA appears to have facilitated some of this; opposition to the disclosure requirements was rebutted on the basis that defence disclosure required little more of defendants than the CJPOA. As Belloni and Hodgson have noted, 'the availability of legal advice continues to be weighed in the scales by the courts, police and policy-makers to justify further encroachment upon the suspect's position'.[64] The next section considers the role of the courts in this attritional process.

55 S Easton, 'Legal Advice, Common Sense and the Right to Silence' (1998) 2(2) E&P 109–22, 119–20.

56 McConville et al (n 49) 131.

57 A Sanders, L Bridges, A Mulvaney and G Crozier, *Advice and Assistance at Police Stations under the 24-hour Duty Solicitor Scheme* (Lord Chancellor's Department 1989) 56.

58 Dixon (n 53) 246.

59 In the largest recent survey, 45% of suspects requested legal advice, with 36.5% actually receiving a consultation: P Pleasence, V Kemp and N J Balmer, 'The Justice Lottery? Police Station Advice 25 Years on From PACE' [2011] Crim LR 3–18. See also L Skinns, who found a 60% request rate, with 48% resulting in a consultation ("'I'm a Detainee; Get Me Out of Here": Predictors of Access to Custodial Legal Advice in Public and Privatized Police Custody Areas in England and Wales' (2009) 49(3) British Journal of Criminology 399–417.

60 S Chakrabarti, 'A Thinning Blue Line? Police Independence and the Rule of Law' (2008) 2(3) Policing 367–74, 369.

61 Leng (n 17) 132.

62 E Cape, 'The Rise (and Fall?) of a Criminal Defence Profession' [2004] Crim LR 50th Anniversary Edition 72–87, 85.

63 Pt II Criminal Procedure and Investigations Act 1996; Criminal Justice Act 2003. Cape adds to this list the limiting of the appeal rights of terrorist suspects by the Anti-Terrorism, Crime and Security Act 2001, the taking of fingerprints, photographs and samples, and some of the amendments to PACE in the Criminal Justice Act 2003 (n 62).

64 F Belloni and J Hodgson, *Criminal Injustice: An Evaluation of the Criminal Justice Process in Britain* (Macmillan 2000) 43.

2 The widening scope of the CJPOA: the Court of Appeal raises the stakes

The CJPOA gave no guidance as to what a proper inference from silence might be but the case-law has developed a presumption that any inference drawn will be adverse.[65] Whilst there have been ebbs and flows in its approach, overall the Court of Appeal has developed the provisions enthusiastically, as happened in both Singapore and Northern Ireland when similar changes were introduced in those jurisdictions.[66] Leng argues that, in its interpretation of s 34, the Court has 'signalled its allegiance to the political objective of the legislation: to establish the norm that suspects should speak in police interview'.[67] The Court now considers 'the police interview and the trial are to be seen as part of a continuous process in which the suspect is engaged from the beginning . . . [a] benign *continuum* from interview to trial'.[68] Whilst directions are given to the jury in each case about the presumption of innocence and the burden of proof, 'to permit the adverse inference must alleviate the prosecution's burden to some degree'.[69]

The Court has made clear that juries should scrutinise both solicitors' reasons for recommending no-comment interviews and suspects' motivations for accepting that advice; in effect, such conduct now has to be justified. Little attention has been given to the effect of these decisions on the ability of defence solicitors to perform their primary function of protecting their clients, and also of ensuring that the police do not exceed their powers. As illustrated below, the judgments are insufficiently consistent for legal representatives to have confidence that advising no-comment interviews will not prejudice their clients should their cases proceed to trial. This section examines the Court's approach to the CJPOA, in particular regarding silence on legal advice, and the impact this has had on legal representatives' ability to fulfil their role.

2.1 The Court's approach to adverse inferences and legal advice

The Crown Court Bench Book, a reference source for judges preparing summings-up for juries, explains that 'the object of section 34 is to deter late fabrication of defences and to encourage early disclosure of genuine defences'.[70] The jury must be directed that it may only draw an inference, if, apart from the defendant's failure to mention facts later relied on in his or her defence, the prosecution case is so strong that it clearly calls for an answer. It must also be satisfied that there is no sensible explanation for the defendant's failure, other than that he or she had no answer at that time or none that would stand up to scrutiny. The jury must be told that an inference should be drawn only if they think it is fair and proper to draw such a conclusion. A defendant must not be convicted wholly or mainly on the strength of the inference.[71] Despite having held previously that the provisions 'should not be construed more widely that the statutory language requires',[72] the Court appears to have gone beyond this. The adverse inference is no longer limited to one of recent invention, but

65 *R v Condron and Condron* [1997] 1 WLR 827; *R v Napper* [1996] Crim LR 591.

66 M Yeo, 'Diminishing the Right to Silence: The Singapore Experience' [1983] Crim LR 88; J Jackson, M Wolfe and K Quinn, *Legislating against Silence: The Northern Ireland Experience* (Northern Ireland Statistics and Research Agency 2000).

67 R Leng, 'Silence Pre-trial, Reasonable Expectations and the Normative Distortion of Fact Finding' 5(4) E&P (2001) 240–56, 249–50.

68 *R v Howell* [2003] EWCA Crim 1; 1 Cr App R 1 [23]–[24].

69 D Hamer, 'The Privilege of Silence and the Persistent Risk of Self-incrimination: Part 1' (2004) 28 Crim LJ 160–78, 161.

70 Crown Court Bench Book <www.judiciary.gov.uk>, 258.

71 Ibid.

72 *R v Bowden* [1999] 1 WLR 823, 827.

now extends to a fact or explanation tailored to fit the prosecution case or which the defendant is thought to have believed would not stand up to scrutiny at the time.[73] When directing jurors how to treat evidence of a no-comment interview, judges should substitute 'conclusions' for 'inferences',[74] which appears to give the 'evidence' greater weight.

The Court has been reluctant to create exclusions to the provisions, preferring to leave the issue to the jury, buttressed by lengthy directions.[75] Redmayne suggests that the 'directions operate rather like a magic formula; so long as they are given by the judge, the jury can be left to draw an inference. The European Court of Human Rights (ECtHR) has taken a similar approach.'[76] The idea that remaining silent on legal advice would protect defendants from adverse inferences[77] was summarily dismissed in the first 'silence' case to reach the Court of Appeal. The Court considered that this 'would render section 34 wholly nugatory'[78] as any competent solicitor would then be bound to advise silence. As with the legislation itself, this 'policy driven'[79] decision ignored how few no-comment interviews there actually were before the law changed, but it resonated with the broader political rhetoric about hardened criminals, abetted by their overly adversarial lawyers, abusing their rights. The jury should be told: 'It is not what the solicitors thought that matters. It is what each defendant thought.'[80] This ignores the fundamental premise of s 58 PACE, which is that the professional offers expert advice to the lay client or, as one defendant retorted, 'Well, what was the point of me having a solicitor there, if I wasn't going to actually take his advice?'[81] The Court explained that it will assume that solicitors will have given proper advice and this must include an explanation of s 34 and the risks of withholding facts from the police. 'It is with the benefit of that advice that an appellant himself makes the decision whether to answer questions or whether to decline to do so and take the risk which goes with it.'[82] The directions fail to consider the generally dismal levels of educational attainment amongst those in custody; the number for whom English is not their first language; the incidence of mental illness; and the extent of drug and alcohol abuse amongst suspects.[83] The ECtHR made little criticism of the provisions but held that the fact that an accused has remained silent on legal advice must be given 'appropriate weight' by the Court, adding that it is essential to a fair trial for the judge to direct the jury not to draw an inference if they are

73 *R v Milford* [2001] Crim LR 330; although in the context of prepared statements, the Court held that the objective of the provisions is early disclosure of a suspect's account and not, separately and distinctly, its subjection to police cross-examination (*R v Knight* [2004] 1 Cr App R 9, CA; *R v Turner (Dwaine)* [2004] 1 Cr App R 24, CA).

74 *Crown Court Bench Book* (n 70) app 2.

75 See the concerns expressed in *R v Bresa* [2005] EWCA Crim 1414, [4]–[5].

76 M Redmayne, 'English Warnings' (2008) 30(3) Cardozo Law Review 1047–88, 1060.

77 RCCJ (n 39) 54; The Law Society's guidelines at the time suggested 'non-tactical' motives would not attract inferences. Home Secretary Michael Howard had, however, stated that this was not his intention as the Bill was debated in Parliament (HC Deb 11 January 1994, vol 235, col 26).

78 *Condron* (n 65) 833.

79 Leng (n 17) 127.

80 *R v Betts; R v Hall* [2001] 2 Cr App R 257, 271.

81 *Howell* (n 68) [6]; Zuckerman (n 32) 30, raised this problem before the CJPOA, referring to judgments from the USA that found such silence to be 'insolubly ambiguous' because it could not be known whether it was due to a sense of guilt, or a preference to exercise the constitutional privilege (*Doyle v Ohio* 426 US 610 (1976) also *US v Hale* 422 US 171 (1975)).

82 *Essa* (n 5) [15].

83 See J J Payne-James, P G Green, N Green, G M C McLachland, M H W M Munro and T C B Moore, 'Healthcare Issues of Detainees in Police Custody in London, UK' (2010) 17(1) Journal of Forensic and Legal Medicine 11–170.

satisfied that the defendant remained silent on legal advice.[84] The domestic courts have 'had some difficulty in giving effect to [this] ruling'.[85] The Court of Appeal appears to have found the crime control[86] arguments more compelling than those of human rights in relation to the right of silence. The Court first held that it is not the *quality* of the suspect's decision to remain silent that matters but the *genuineness* of that decision;[87] it then held that the jury should apply a more stringent objective[88] test in deciding this. To date its parsing of the judgment in this way has not been challenged in Strasbourg.

The Court of Appeal has taken an increasingly elastic view as to when the police are allowed to question a suspect against whom they have sufficient evidence to charge. Section 37(2) of PACE requires custody officers to charge or release a suspect if they determine that they have sufficient evidence to charge them. Initially it was held that inferences should not be drawn in cases where the interviewing officer believed, prior to an interview, that there was sufficient evidence for a successful prosecution.[89] The Court then resiled from this, holding that decisions about prosecution must involve consideration of any explanation given by the suspect.[90] In the 'pushmi-pullyu' way in which government and the judiciary have dealt with these cases, the PACE Codes of Practice were then revised to give effect to this. This gives greater scope to the interviewing officer to continue questioning and in a manner which is difficult to reconcile with s 37(2).[91]

The Court reinforced its commitment to establishing the norm of suspects answering questions in *Hoare*:

> It is not the purpose of section 34 of the 1994 Act to exclude a jury from drawing an adverse inference against a defendant because he genuinely or reasonably believes that, regardless of his guilt or innocence, he is entitled to take advantage of that advice to impede the prosecution case against him. In such a case the advice is not *truly* the reason for not mentioning the facts . . . Legal entitlement is one thing. An accused's reason for exercising it is another.[92]

It is unclear how it can be 'unreasonable' for a suspect to follow legal advice, other than by the implication that it is being used as a pretext to evade conviction. As for discerning the true explanation, this is an entirely speculative exercise, as the jury cannot know the defendant's reasoning; there may not even be a 'true' explanation. Despite the caveat about innocence, the Court's reasoning makes sense only if the defendant is guilty. A factually innocent suspect may be affronted at being arrested and refuse to answer police questions in protest; this would truly be the reason for remaining silent but not one for which the Court allows in the above judgment. Although jurors must be directed not to convict the defendant wholly or mainly on the strength of the inference, if they decide the inference is

84 *Condron v UK* (2001) 31 EHHR 1, [60]–[62].

85 Lord Phillips of Worth Matravers, 'Trusting the Jury' (Criminal Bar Association Kalisher Lecture 2007) <www.judiciary.gov.uk/Resources/JCO/Documents/Speeches/lcj_trusting_juries_231007.pdf> 6 accessed 27 June 2013.

86 R Pattenden, 'Silence: Lord Taylor's Legacy' (1998) 2(3) E&P 141–65, 164.

87 *Betts and Hall* (n 80) 270.

88 *R v Hoare and Pierce* [2005] 1 WLR 1804; [2004] EWCA Crim 784, following *Howell* (n 68); and *R v Knight* [2003] EWCA Crim 1977, rather than the subjective test preferred in *Betts and Hall* (n 80).

89 Para 11.4 of the old Code C (see E Cape, 'Detention without Charge: What Does "Sufficient Evidence to Charge" Mean?' [1999] Crim LR 87); *R v Pointer* [1997] Crim LR 676; and *R v Gayle* [1999] Crim LR 502.

90 *R v McGuinness* [1999] Crim LR 318; *R v Ioannou* [1999] Crim LR 586; *R v Elliott* [2002] EWCA Crim 931; *Howell* (n 68).

91 Para 11.6 of the new Code C. See E Cape 'The Revised PACE Codes of Practice: A Further Step Towards Inquisitorialism' [2003] Crim LR 355–69.

92 *Hoare and Pierce* (n 88) [54]; see also *Howell* (n 68) [23]–[24].

adverse, this now seems likely to be decisive. A jury might accept that a suspect remained silent on legal advice, yet convict upon other evidence; it is unlikely to assume a suspect accepted the lawyer's advice because he or she was guilty and then acquit.[93] This has worrying consequences for the value of legal advice:

> The value of legal advice is thus diminished. If defendants can never be sure that they are acting reasonably in relying on the advice of their lawyer, then they can never be sure that they should accept their lawyer's advice. If they cannot be sure about that, then it raises the fundamental question of the utility of legal advice at the police station. Custodial legal advice may be guaranteed by the European Convention on Human Rights, but its value as such is in danger of being seriously eroded.[94]

2.2 The implications for legal representatives

As well as making it more difficult for the accused to follow the advice of their lawyers, the judgments of the courts in these cases have made it much harder for legal representatives to decide how best to advise their clients. The Court of Appeal has stated that the kind of circumstance which may *justify*[95] a solicitor advising silence (an interesting choice of words) will be matters such as the suspect's condition (ill-health, in particular mental disability; confusion, intoxication or shock), or genuine inability to recollect events without reference to documents, or communicating with others. In deciding what is reasonable, a court is to consider the particular accused with such qualities, knowledge, apprehensions and advice as he or she was shown to have had at the time.[96] This has been interpreted inconsistently, however. The Court rejected the decision of the Condrons' solicitor that, contrary to the view of the police doctor, they were unfit to be interviewed, due to the effects of heroin withdrawal.[97] It is difficult to conceive of a situation in which it would be more reasonable for a suspect to follow the advice of his representative than where, as in *Roble*,[98] the solicitor considered his understanding of English to be insufficient to deal with difficult legal concepts and she was unclear what his instructions were. The Court of Appeal disagreed, however; as the solicitor did not give evidence of all the facts her client had told her, it did not know the basis upon which she had formulated her advice. To follow the judges' own reasoning, if they are not considering the correctness of the solicitor's advice, they do not need to know the logic underpinning it.[99] Such analysis demonstrates further the scepticism of the courts to those making no comment. Reasons for the legal representative's advice are deemed irrelevant when they are adduced to avoid inferences or to support excluding the interview (it is the defendant's reasons for silence that are significant) but the legal representative's lack of reasons may be used as justification for drawing inferences.

The Court refused from the outset to exclude no-comment interviews where the legal advice was tactically based, on the grounds that the advice may not be bona fide; there may be more than one reason for giving it; and 'it is not so much the advice given by the solicitor, as the reason why the defendant chose not to answer questions that is important and this is

93 What Wolchover describes as 'classic *petition principia*—assuming what is to be proved as an essential component of the would-be proof': D Wolchover, 'Serving and Saving Silent Suspects: Part 2' (12 February 2011) 175 Criminal Law and Justice Weekly 86–88, 88.

94 Cape (n 21) 402.

95 *Howell* (n 68) [24].

96 *R v Argent* [1997] 2 Cr App R 27.

97 *Condron* (n 65).

98 *R v Roble* [1997] Crim LR 449.

99 See Cape (n 21) 402.

a question of fact which may be very much in issue'.[100] The 'acid question' in these cases is not was it reasonable to rely on the solicitor's advice, but rather, could the appellant reasonably have been expected to say what is then relied upon at trial?[101] The Court went on to give examples of what would not be considered reasonable grounds for making no comment: the absence of a written statement from the complainant; the likelihood that the complainant will withdraw; or the solicitor's belief that the suspect will be charged anyway.[102] These are all 'tactical' reasons that would give practical effect to the principle that the burden of proof lies on the prosecution. These judgments suggest a worryingly instrumental view that infringing suspects' rights matters less than gaining a confession if it convicts the guilty – a viewpoint which PACE had had an significant role in changing – and assumes that little harm will come to the innocent who should be willing to submit to the indignity of an illegal interview in order to exculpate themselves. The Court has reaffirmed this view of the primacy of co-operation, leaving inferences to the jury in a case in which silence was advised due to the expiration of the lawful detention period for one suspect and it was not clear that the police had evidence to suggest that the other had committed an offence.[103]

The extraordinary judgment in *Howell* shows how far the Court has adopted the crime-control rhetoric that infused the CJPOA. It makes clear that, regardless of the presumption of innocence, suspects are expected to co-operate with proceedings against themselves and has extended the effect of the Act much further than had ever been anticipated.

> . . . if currency is given to the belief that if a suspect remains silent on legal advice he may systematically avoid adverse comment at his trial. And it may encourage solicitors to advise silence for *other than good objective reasons* . . . There must always be soundly based objective reasons for silence, sufficiently cogent and telling to weigh in the balance against the clear public interest in an account being given by the suspect to the police. Solicitors bearing the important responsibility of giving advice to suspects at police stations must always have this in mind.[104]

This is a fundamental change in the nature of the role of the legal adviser for which there is no legislative authority; in fact the PACE Codes appear to contradict this directly:

> The solicitor's *only* role in the police station is to protect and advance the legal rights of their client. On occasions this may require the solicitor to give advice which has the effect of the client avoiding giving evidence which strengthens a prosecution case. The solicitor may intervene in order to seek clarification, challenge an improper question to their client or the manner in which it is put, *advise their client not to reply to particular questions*, or if they wish to give their client further legal advice.[105]

Defence solicitors now face a conflict between their statutory and professional obligations,[106] to advise and to act in their clients' best interests, and the changed approach of the Court. This impasse appears impossible to resolve in their clients' favour whilst the Court maintains its current stance.

---

100 *Condron* (n 65) 833.

101 *Essa* (n 5) [15].

102 Howell (n 68) [24].

103 *Hoare and Pierce* (n 88) [3].

104 *Howell* (n 68) [23]–[24], emphasis added.

105 PACE Code C, para 6D, emphasis added.

106 'Advice to Practitioners from the Criminal Law Committee of the Law Society', *Criminal Practitioners Newsletter* (October 1994); 'Law Society Advice to Practitioners', *Criminal Practitioners Newsletter* (July 1997); 'Changes in the Law Relating to Silence', *Criminal Practitioners Newsletter* (January 2006).

Cape notes the 'signs of a growing antipathy towards adversarial principles and the adversarial role of defence lawyers'.[107] It has been suggested, both in the Court of Appeal and ECtHR, that a good reason for not drawing an adverse inference would be 'bona fide advice received from his lawyer'.[108] This is a troubling term as it suggests by implication that lawyers may advise in bad faith and, following the logic of the case-law described above, this could include advising suspects to make no comment for tactical reasons. Such a view assumes the public interest inheres only in achieving convictions, rather than in cases being investigated and tried in an appropriate manner. The term was used in *Averill v UK*,[109] a case that related to the analogous provisions operating in Northern Ireland at a time when 'the security forces in Northern Ireland violate[d] basic international standards safeguarding the role of lawyers on a regular basis'.[110] Defence solicitors are officers of the court[111] and, in the absence of evidence to the contrary, should be presumed to be acting in good faith. Lawyers 'should never be identified with their clients or their clients' causes as a result of discharging their functions'.[112]

The old legal aphorism of 'if in doubt, say nowt [nothing]', whilst never satisfactory, or as widespread as the police believed, became positively negligent after April 1995. Legal representatives have to make decisions quickly and in difficult circumstances, often in ignorance of some or all of the facts, knowing that their decision could have negative repercussions for their clients:

> The imposition on the lawyer is really quite tough, because we are now in a very unenviable position of, if our advice is wrong and we don't give adequate legal protection to our client . . . it's an element that can be used to convict them . . . There is more thought process required.[113]
>
> At 3 o'clock in the morning you've been dragged out of bed, your pyjamas are sticking out from underneath your jeans and you're trying to write a cogent note about why you've given certain advice which somebody is going to sniff over nine months later.[114]

Poor legal advice is not, of itself, a ground of appeal, rather the appellant has to show that the incompetence led to identifiable errors or irregularities in the trial, which themselves rendered the process unfair or unsafe.[115] If the Court of Appeal tends to the view that solicitors should advise their clients to answer questions, it seems unlikely that it will sympathise with an appellant who claims to have been ill advised to answer questions. If it

107 Cape (n 62) 85–86. He cites *Gleeson* [2003] EWCA Crim 3537 and *Hughes v DPP* [2003] EWHC 2470 (Admin). See also R *(on the application of the DPP) v Chorley Justices* [2006] EWHC 1795; *Arthur JS Hall and Co v Simons (AP)* [2000] UKHL 38, [2002] 1 AC 615).

108 *Averill v UK* (2001) 31 EHRR 36, [47]; *Condron* (n 65) 833.

109 *Averill v UK* (2001) 31 EHRR 36, [47].

110 Lawyers Committee for Human Rights, *Human Rights and Legal Defense in Northern Ireland: The Intimidation of Defense Lawyers and the Murder of Patrick Finucane* (Lawyers Committee for Human Rights 1993) 25. See also: Sir J Stevens, *Overview and Recommendations*, 17 April 2003 (PSNI 2003); U Lucy, 'Improper Inference: The Perils of Defending Suspected Terrorists in Northern Ireland' (2003) 15(2) Pace International Law Review 411–36; K McEvoy, 'What Did the Lawyers Do During the "War"? Neutrality, Conflict and the Culture of Quietism' (2011) 74 MLR 350–84.

111 See Solicitors' Regulation Authority, *Code of Conduct* (2011) <www.sra.org.uk/solicitors/handbook/code/content.page> accessed 27 June 2013.

112 Principle 18 of United Nations, *Basic Principles on the Role of Lawyers*, 7 September 1990 <www.unhcr.org/refworld/docid/3ddb9f034.html> accessed 27 June 2013.

113 Legal Executive B1/ii.

114 Solicitor D9/ii.

115 *R v Day* [2003] EWCA Crim 1060. See E Cape 'Incompetent Police Station Advice and the Exclusion of Evidence' [2002] Crim LR 471–84.

takes the view that suspects should volunteer information to the police, it is unlikely to accept poor advice to make no comment as a ground of appeal.

The courts have expressed their expectation that suspects will co-operate with police questioning. There is, however, no reciprocal obligation on the police to disclose their case to suspects or their advisers. Contrary to early expectations, and indications from the Court of Appeal,[116] inadequate police disclosure does not justify the exclusion of a no-comment interview; rather it should be left to the jury to consider the reasonableness of the advice.[117] Because representatives are unsure of the risks of advising no comment if the police refuse to disclose their case, some said that they may advise the client to 'say something and stop the interview if you are not too sure',[118] or say that they will intervene if the police introduce undisclosed material. There is a danger, of which officers are well aware, that the suspect may have answered before the adviser can respond:[119]

> We were having a lot of interviews stopped, 'I need a consultation with my client' and to be honest, I'd rather do that, I'd rather go into an interview and say something and the solicitor knows nothing about it. I mean, it's a particularly selfish way of looking at it, I suppose the suspect, nine times out of ten has answered your question before the solicitor says 'Oh, hang on a minute.'[120]

It cannot be fair that evidential significance attaches to the suspect's decision whether or not to reply in interview, when this decision must be made without full knowledge of the police case. Without disclosure, there can be no balance in the interview, as legal representatives have no means of checking the veracity of the police account, or of determining the lawfulness of the arrest, detention and questioning. A no-comment interview was one of the few sanctions or negotiating ploys available to legal representatives if they thought that the police were on a 'fishing expedition', had acted improperly, or were refusing to disclose their case. The CJPOA restricted this means of negotiation, which has circumscribed further the benefits of legal advice.[121] Some solicitors have tried to avoid inferences by submitting a prepared statement before a no-comment interview. If this gives full details of the defence then no inference should be drawn.[122] The Court has warned that inferences may be drawn if a more detailed or inconsistent account is given at trial[123] and cautioned that this 'may prove to be a dangerous course for an innocent person who subsequently discovers at the trial that something significant has been omitted'.[124]

Bucke et al[125] found a drop in the number of suspects refusing to answer all questions from 10 per cent before the CJPOA to 6 per cent afterwards. The Court of Appeal has shown little sympathy for the difficulties the CJPOA presents for legal advisers and has failed to recognise that the safeguards provided at trial, such as allowing inferences to be drawn only once the prosecution has established a prima facie case, do not apply when the suspect and solicitor actually make their decisions. The ECtHR has sought gently to

116 *Roble* (n 98). See R J Toney, 'Disclosure of Evidence and Legal Assistance at Custodial Interrogation: What Does the European Convention on Human Rights Require?' (2001) 5 E&P 39, 53–54.

117 *Argent* (n 96); *R v Imran and Hussein* [1997] Crim LR 754.

118 Solicitor D13/ii.

119 Detective Constable A5.

120 Detective Constable A3.

121 Jackson (n 38) 160.

122 See Crown Court Bench Book (n 74).

123 *R v Knight* [2004] 1 Cr App R 9, CA; followed in *R v Turner (Dwaine)* [2004] 1 Cr App R 24, CA.

124 Turner (n 123). See also *R v Faisal Khan Mohammad* [2009] EWCA Crim 1871.

125 Bucke et al (n 13).

reinforce the 'paramount importance'[126] of legal advice in ensuring fairness to suspects at the police station, acknowledging that the provisions place the accused in a 'fundamental dilemma'.[127] This led to the CJPOA being amended so that inferences can no longer be drawn from a suspect's failure to answer questions in interview if legal representation has been denied.[128] What both the ECtHR and Court of Appeal have failed to recognise, however, is the fundamental dilemma that the *legal representative* now faces. Advisers now have to assess not only the current strength of the police case, whether or not the police are willing to disclose it, but also predict its potential strength should the case come to trial. Advising suspects to answer questions may provide the missing information that allows them to be charged; whereas recommending no comment may result in inferences being drawn that strengthen a weak prosecution case to the required standard of proof should the case get to trial. The risk of invoking inferences has diminished the protective function of the lawyer and undermined their bargaining power to ensure the fair and legal treatment of clients at the police station. This inhibits the defence from testing the police case and in practice may allow the police to question the suspect on dubious legal authority. It appears that the Court now considers the role of the adviser should be to facilitate the expeditious processing of their clients; an advantage for the legal system, but not necessarily in their clients' best interests, and an interpretation of the CJPOA which goes far beyond that anticipated when it was enacted.[129]

3 A 'fundamental dilemma' v a 'fundamental condition': how inferences trumped legal professional privilege

Part one of this article suggested how the CJPOA led to further legislative encroachments on defendants' rights; this section explores how the Court of Appeal has allowed the legislation to encroach upon legal professional privilege, a principle it had hitherto defended robustly. Communications between lawyer and client have a special legal status, preserving their confidentiality in order to encourage full and uninhibited disclosure.[130] Once established, no exception should be allowed to the absolute nature of this privilege.[131] In a ringing endorsement, the House of Lords described legal professional privilege as a fundamental right, protected by the ECHR and held:

> . . . a man must be able to consult his lawyer in confidence, since otherwise he might hold back half the truth. The client must be sure that what he tells his lawyer in confidence will never be revealed without his consent. Legal professional privilege is thus much more than an ordinary rule of evidence, limited in its application to the facts of a particular case. *It is a fundamental condition on which the administration of justice as a whole rests.*[132]

The leading case on privilege of *Derby Magistrates' Court*[133] involved a request for privileged information that might yield potentially exculpatory evidence in a murder trial.

126 *Murray (John) v UK* (1996) 22 EHRR 29, 67.

127 Ibid.

128 S 58 Youth Justice and Criminal Evidence Act 1999. If access to legal advice is denied, then suspects should be given the pre-CJPOA caution (PACE Code of Practice C, Annex C).

129 S Cooper, 'Legal Advice and Pre-trial Silence: Unreasonable Developments' (2006) 10(1) E&P 60–69, 69.

130 *Campbell v UK* (1993) 15 EHRR 137. Legal privilege is defined in s 10 of PACE, which gives effect to the common law position (*R v Central Criminal Court ex p Francis & Francis* [1989] AC 346, 392). See J Auburn, *Legal Professional Privilege: Law and Theory* (Hart Publishing 2000).

131 *R v Derby Magistrates' Court ex p B* [1996] AC 487, HL; and *B v Auckland District Law Society* [2003] 2 AC 736, PC.

132 *Derby Magistrates' Court* (n 131) 507, emphasis added. For a critical review of this decision, see C Tapper, 'Prosecution and Privilege' (1996) 1(5) E&P 20.

133 *Derby Magistrates' Court* (n 131).

Notwithstanding the importance of this claim, protection of privilege was regarded as 'so fundamental that it "trumps" . . . any claim of a defendant to show his innocence by relying on material protected by another's privilege'.[134] Whilst acknowledging that the right was not absolute, their Lordships declined to undertake a balancing exercise, reasoning that 'once any exception to the general rule is allowed, the client's confidence is necessarily lost'. The Court noted that Parliament had not addressed privilege and thus refused to erode it.[135] In other cases it has held that the right to legal confidentiality cannot be overridden by general or ambiguous statutory words; an intention to override must be stated expressly or appear by necessary implication.[136] In the CJPOA cases, judges have been markedly less robust in their defence of privilege.[137]

Many commentators assumed that legal advice would be a reasonable excuse for failure to mention facts at the police station.[138] The Court thus had to reconcile preserving this fundamental condition with the risk of 'guilty' suspects exploiting legal privilege to avoid inferences from their silence. The manner in which it has done so appears rooted firmly in the post-CJPOA expectation that silence has to be justified, and is in contrast to its previous position.[139] As discussed above, the Court gave the submission short shrift, fearing a 'coach and horses'[140] would be driven through the provisions if this were allowed. A simple assertion of reliance on legal advice, without explanation, will be unlikely to avert inferences. Defendants who remain silent following legal advice now have to balance the risks of waiving privilege against the potential advantages of calling their adviser to explain the basis of their advice, or to rebut allegations of subsequent fabrication. This interpretation of the provisions has been held not to contravene Article 6 ECHR as the suspect experiences only the 'indirect compulsion' to waive privilege of being able to offer merely a bare explanation as a defence.[141] In practice this amounts to little more than Hobson's choice, as suspects either have an unexplained silence from which inferences may be drawn, or they seek to explain this and have to waive privilege. Whilst the Court is concerned that the defendant should not 'have his cake and eat it',[142] it would appear that the prosecution can and does enjoy the benefits of being able either to use potentially damaging disclosures from the suspect, or of inviting inferences from a suspect's silence.

The Court of Appeal sought to clarify the law on privilege in s 34 cases in *Seaton*.[143] Having reviewed the relevant authorities,[144] the Court reiterated that legal professional

134 *R v Loizou* [2006] EWCA Crim 1719, [40].

135 *Derby Magistrates' Court* (n 131).

136 *R (Morgan Grenfell & Co. Ltd) v Special Commr of Income Tax* [2003] 1 AC 563, HL; and *B v Auckland District Law Society* (n 131).

137 The Court went on to interpret the Regulation of Investigatory Powers Act 2000, which contains no express provision about privilege, to permit the covert surveillance of meetings between defendants and their lawyers, without express Parliamentary authority (*Re McE* [2009] 1 AC 908). See Bar Council 'Bar Council: Private Legal Advice Must Be Protected', press release, 2 February 2012 <www.barcouncil.org.uk/media-centre/news-and-press-releases/2012/february/bar-council-private-legal-advice-must-be-protected> accessed 27 June 2013.

138 Dennis (n 43) 17; R Munday, 'Inferences from Silence and European Human Rights Law' [1996] Crim LR 370–85.

139 For a more optimistic interpretation see D Wolchover, 'Serving Silent Suspects' pts 1–3 (2011) 175 Criminal Law and Justice Weekly 71–72, 86–88, 104–06.

140 *R v Beckles* [2004] EWCA Crim 2766; [2005] 1 WLR 2829, [43].

141 *Condron v UK* (n 84) 22.

142 *Bowden* (n 72) 831.

143 *R v Seaton* [2010] EWCA Crim 1980, [43].

144 *R v Wilmot* (1989) 89 Cr App R 341; *Condron* (n 65); *Bowden* (n 72); *R v Wishart* [2005] EWCA Crim 1337; and *R v Loizou* [2006] EWCA Crim 1719.

privilege is of paramount importance and that there is no question of balancing privilege against other considerations of public interest. In the absence of waiver, no question can be asked which intrudes upon privilege. The mere assertion by the accused that the reason for not answering questions was on legal advice does not amount to waiver. Yet its judgment was not wholly consistent. Of greatest significance was the Court's finding that the previous cases had misinterpreted *Wilmot*.[145] The Court held that a defendant who adduces evidence of the content of, or reasons for, such advice, beyond the mere fact of it, does waive privilege, at least to the extent of allowing cross-examination about what the solicitor was told and whether this can be the true explanation for the suspect's silence.[146] Once this happens, the legal representative can be asked about any other reasons for the advice, including whether it was for tactical reasons, the nature of the advice, and the factual premise upon which it was based.[147] Where a defendant does not waive privilege by calling the solicitor to give evidence, the judge may direct the jury:

> You have no explanation for the advice in this case. It is the defendant's right not to reveal the contents of any advice from his solicitor or what transpired between himself and his solicitor. At the same time he has a choice whether to reveal that advice and thereby reveal all that transpired between himself and his solicitor.[148]

If the comment is fair, another party can comment upon the fact that the solicitor has not been called to confirm something which, if true, could be confirmed easily. This does not waive privilege entirely; the test is fairness, or the avoidance of a misleading impression, or both.

Some of the cases relating to legal privilege have turned on quite technical legal questions regarding waiver. What has not been addressed by the courts is the potentially deleterious effects of using the defendant's legal representative, the person supposed to 'protect and advance the legal rights of their client' potentially to discredit them. The Court in *Seaton* raised the possibility of the Crown being able to call the defence solicitor to give evidence on these matters, albeit limited to the point in issue. It noted the danger of the solicitor 'trespassing inadvertently beyond the part of the communications which the defendant had opened up' and that 'in reality that would not be practicable', but focused on the logistics of such an exercise rather than the potential ramifications of such an action. Even if the cross-examination were confined to whether or not the defendant had told the solicitor the relevant fact that had not been revealed in interview, it would seem likely that defendants might lose confidence in the confidentiality of proceedings and this might affect the frankness in future cases of the 75 per cent who are repeat offenders.[149]

The possibility of the legal representative having to give evidence has introduced tensions into the solicitor–client relationship as advisers now consider the ramifications for themselves as well as, or perhaps instead of, their clients.[150] There was a deepening concern that 'a barrier will be created between suspect and adviser: the whole purpose of independent legal advice and the basic principles of adversarialism will be undermined'.[151] Suspects should sign the advisers' attendance notes to confirm they have made the decision to remain silent but, even before the courts required this,[152] in order to protect themselves

145 *Wilmot* (n 144).

146 *R v Seaton* [2010] EWCA Crim 1980, approving *Bowden* (n 72).

147 *Bowden* (n 72).

148 *R v Bresa* [2005] EWCA Crim 1414, [49].

149 Criminal Justice Statistics Quarterly Update to September 2011 Ministry of Justice Statistics Bulletin 2012.

150 Cape (n 21).

151 Leng (n 17) 127.

152 *Criminal Practitioners Newsletter* (2006) (n 106).

from criticism at trial, some firms had produced all-encompassing disclaimers for their clients to sign:

> I [name] have been advised by my Solicitors, [name and address of firm], that an adverse inference may be drawn if I make no comment in an interview. Nevertheless I wish to have a 'no comment' interview and fully understand the consequences that may subsequently arise. Signed ___ Name (please print) ____ Date ___[153]

Some solicitors appear more concerned with protecting their own positions than acting solely in their clients' interests:

> They're thinking I'm there to advise them. I say 'it's got to be your decision at the end of the day'. I don't want to get the blame . . . I'm fearful of that. That's why I try to put the guilt trip on the client as it were. I let him decide. More often than not, I get him to sign something.[154]

> It's so rare that you advise them to remain silent, that you make sure you've got concrete reasons for doing it, because you always think 'I may have to take the stand about this one day.'[155]

Few legal representatives interviewed had given evidence as to why they had advised a no-comment interview. A minority were concerned about the prospect, through fear ('It scared the living daylights out of me');[156] a belief that solicitors should not have to justify their advice;[157] or because of the financial implications of losing a client.[158] (A solicitor cannot represent a client and be called as a witness in their cause, so, in the magistrates' court, new representation would have to be sought which is not in the solicitor's interests.) Barristers may also experience conflict from the commercial pressure not to alienate the solicitors who give them work:

> It's such a tight community . . . one day you can be prosecuting, you know [X]'s client and the next day, you're acting on his behalf, and the day before you called him an absolute wazzock! [idiot][159]

Whilst the Court of Appeal has observed the niceties about legal privilege, the change in approach it has developed through its interpretation of the CJPOA provisions has facilitated incursions into another fundamental protection for defendants. Defence solicitors having to testify about the advice they gave, in however restricted a form, allows the prosecution to probe further the 'genuineness' and 'reasonableness' of the suspect's decision to make no comment. It is not known what effect it may have on the jury of seeing a defence representative cross-examined, but it seems likely to attract a disproportionate amount of the jury's attention compared to its evidential value.[160] In a most literal way, defence solicitors may now have to justify under oath their decision to advise no comment.

4 Conclusion

The rhetoric around the CJPOA advocated rebalancing a system that was described – rather than demonstrated – as being too favourable to criminals. As well as facilitating further

153 Firm J14.
154 Solicitor D13/ii.
155 Solicitor A13/iii.
156 Solicitor A12/ii.
157 Solicitor A13/iii.
158 Solicitor B2/i.
159 Barrister B1.
160 Birch (n 16) 786.

legislative encroachments by Parliament, it appears to have changed the culture in which defendants are tried. Many of the changes were criticised as being inappropriate to an adversarial system, leaving the accused with the worst of both worlds: being investigated in a semi-inquisitorial manner, the results of which are then deployed in an adversarial context.[161]

The CJPOA was founded on the contention that 'criminals,' abetted willingly or witlessly by their lawyers, were exploiting the right of silence to evade justice. Such a view paid scant regard to the research or to the consequences for the presumption of innocence and the burden of proof. The protections offered to suspects in custody by PACE are undoubtedly an improvement on the common law but the suggestion that the right to legal representation gave suspects an unfair advantage over the police was demonstrably flawed. Only a minority of suspects was legally represented and the quality of that representation was variable. The police are gatekeepers of suspects' rights and can circumvent or minimise the benefits of many of these protections. The refusal to answer police questions was hitherto the only immediate sanction that suspects and legal advisers had in response to the police either abusing their powers or acting beyond them; this now carries evidential risks. More suspects now have access to legal advice but their representatives are able to do less. The protective benefit of legal advice, a fundamental requirement of a fair trial, is devalued by the quandary in which legal representatives are now placed of having to choose between the risks of allowing suspects to answer questions and potentially inculpate themselves; or of facing adverse inferences, should prosecutions follow. Whilst the direct effects of the Act have been limited by the relatively small numbers of suspects making no comment, its effects upon legal representatives have shifted the balance of power in custodial interrogation further in favour of the police.

The Court of Appeal appears to have absorbed the crime-control assumptions underlying the CJPOA. In its interpretation of the Act, it has shown increasing impatience with the idea that suspects should not have to facilitate the prosecutorial process and has disapproved of solicitors representing their clients in an adversarial manner. It has refused most attempts to circumscribe the effects of the CJPOA, even when encouraged to do so by the ECtHR. It has shown little understanding of the difficulties that solicitors face in their work and has sought to restrict their role in a manner which arguably contravenes the intentions of Parliament. The Court of Appeal has made some attempt to limit the extent of the waiver of privilege, but solicitors are now acutely aware that they may have to give evidence about their advice, thus losing the client if the case is heard summarily, and they fear damaging their reputation with the client if their advice is deemed to be wrong, and risk damaging their client's case if their testimony is not well-received. The Court of Appeal has pursued the idea that no-comment interviews are not a legitimate tactic in an adversarial situation, but rather are unnecessarily obstructive of justice. Jackson observed that the CJPOA provisions have made police station questioning part of the trial process;[162] in its interpretation of the Act, the Court of Appeal has gone further and imported the traditional suspicions that police officers have of lawyers into the courtroom to the detriment of defendants, legal advisers and to two fundamental principles of justice. Twenty years later the right of silence is little discussed but its effects continue to reverberate.

161 Cape (n 115).

162 Jackson (n 22).

NILQ 64(4): 485–505

Mental (in)capacity or legal capacity? A human rights analysis of the proposed fusion of mental health and mental capacity law in Northern Ireland

EILIONÓIR FLYNN

National University of Ireland, Galway

Introduction

The United Nations Convention on the Rights of Persons with Disabilities (CRPD) entered into force in 2007 and sets out in Article 12 that all human persons, regardless of their decision-making capabilities, should enjoy 'legal capacity' on an equal basis – that is, the right to be recognised as a person before the law, and the subsequent right to have one's decisions legally recognised.[1] Simultaneously, the Bamford Review of Mental Health and Learning Disability in Northern Ireland published its final report that year, setting out principles for future legislation 'to provide equally for all circumstances in which an individual's autonomy might be compromised on health grounds',[2] stating that such legislation should value the rights of persons with mental health needs or a learning disability, including 'their rights to full citizenship, equality of opportunity, and self determination'.[3]

The continuities and discontinuities between the subsequent legislative proposal for the Mental Capacity (Health, Welfare and Finance) Bill (Northern Ireland) and the emerging consensus in international human rights law on the need for substitute decision-making regimes to be replaced by supported-decision-making[4] have not been fully examined to date. In this paper, I seek to address the tensions between these two positions, in light of the Concluding Observations of the Committee on the Rights of Persons with Disabilities on Article 12 CRPD. Although mental health laws which permit involuntary detention and

1 See, for example, Tina Minkowitz, 'The United Nations Convention on the Rights of Persons with Disabilities and the Right to Be Free from Nonconsensual Psychiatric Interventions' (2006–2007) 34 Syracuse Journal of International Law and Commerce 405, 408; Amita Dhanda, 'Legal Capacity in the Disability Rights Convention: Stranglehold of the Past or Lodestar for the Future?' (2006–2007) 34 Syracuse Journal of International Law and Commerce 429; G Quinn with A Arstein-Kerslake, 'Restoring the "Human" in "Human Rights": Personhood and Doctrinal Innovation in the UN Disability Convention' in C Gearty and C Douzinas (eds), *The Cambridge Companion to Human Rights Law* (Cambridge University Press 2012)

2 Bamford Review of Mental Health and Learning Disability (Northern Ireland), *A Comprehensive Legislative Framework* (Belfast 2007) 3.

3 Ibid.

4 See, for example, *Consideration of Reports Submitted by States Parties under Article 35 of the Convention: Concluding Observations, Tunisia*, Committee on the Rights of Persons with Disabilities (CRPD), 5th Session, 4, UN Doc CRPD/C/TUN/CO/1 (11–15 April 2011); *Consideration of Reports Submitted by States Parties under Article 35 of the Convention: Concluding Observations, Spain*, Committee on the Rights of Persons with Disabilities (CRPD), 6th Session, 5, UN Doc CRPD/C/ESP/CO/1 (19–23 September 2011).

treatment of persons with psycho-social disabilities are also in need of reform according to Article 14 CRPD, I will focus my analysis in this paper on the need for reform in light of Article 12, since the proposed framework for Northern Ireland intends to apply the same legal standards to persons with mental health issues as to those with any other perceived deficits in decision-making ability. I will also highlight comparative law reform developments on legal capacity in other jurisdictions since the entry into force of the CRPD, with a particular focus on developments in the Republic of Ireland, given the requirement in the Good Friday/Belfast Agreement to ensure equivalence of protection for human rights on both sides of the border.[5]

Prior to the discussion of the law which follows, it is important to clarify at the outset the terminology used in this paper. The term 'mental capacity' refers to the decision-making skills of an individual, which naturally vary among individuals and may be different for a given individual depending on multiple factors (including environmental, social and other factors). By contrast, legal capacity is the recognition of an individual as both a holder of rights (with legal status or personality) as well as an actor in law (legal agency).[6] In modern times, mental capacity has been used as a proxy for legal capacity in many instances – meaning that, if a person does not meet a certain standard of decision-making ability, legal recognition of the validity of that individual's decision will be removed from the person and vested in a third party, who generally makes a decision based on the perceived 'best interests' of the person. This removal of legal capacity (generally based on a functional assessment of mental capacity) is referred to as substitute decision-making and can take many forms (including adult guardianship, and best interests decision-making under the Mental Capacity Act 2005 in England and Wales).

With the entry into force of the CRPD in 2007, the prevalence of the functional approach to mental capacity assessment, and the subsequent removal of legal capacity from individuals, began to be challenged from a human rights perspective. As will be discussed in further detail below, the CRPD introduced what Arstein-Kerslake and I refer to as the support model of legal capacity[7] – based on the notion that legal capacity is a universal attribute and not a privilege to be granted only to individuals who have certain decision-making skills. This does not mean that we do not acknowledge that individuals have different levels of decision-making ability, or 'mental capacity', rather, we suggest that, instead of responding to deficits in decision-making by removing the legal authority to make binding decisions from individuals, the response most in keeping with the values and principles of the CRPD is to offer support to individuals in exercising their legal capacity. This support can take many forms, described further below, but one of the most commonly referred to is 'supported decision-making', which can be described as a formal decision-making agreement between an individual and one or more supporters, setting out areas of decision-making with which the person would like assistance, and which must be respected

5 C Ó Cinnéide, *Equivalence in Promoting Equality: The Implications of the Multi-party Agreement for the Further Development of Equality Measures for Northern Ireland and Ireland* (Equality Authority 2005) 1–64 <www.equality.ie/en/Publications/Policy-Publications/Equivalence-in-Promoting-Equality.html> accessed 11 October 2012.

6 International Disability Alliance, 'Legal Opinion on Article 12 of the CRPD' (21 June 2008) 1 <www.internationaldisabilityalliance.org/representation/legal- capacity-working-group> accessed 13 October 2011; Office of the United Nations High Commissioner of Human Rights, *Legal Capacity* (Background Conference Document 2005) 1, 13 <www.un.org/esa/socdev/enable/rights/ahc6documents.htm> accessed 13 October 2011.

7 E Flynn and A Arstein-Kerslake, 'Legislating Personhood: Realising the Right to Support in Exercising Legal Capacity' (2013) 9(4) International Journal of Law in Context (forthcoming).

by third parties. Having clarified the terms to be used in the rest of the discussion, I will now turn to the current law in Northern Ireland and proposals for reform.

Mental capacity and mental health: current law in Northern Ireland

The main legislative provision governing mental health in Northern Ireland is the Mental Health (Northern Ireland) Order 1986. This order provides for the involuntary detention and treatment of persons with 'mental disorders'. 'Mental disorder' is defined in the Order as 'mental illness, mental handicap and any other disorder or disability of mind'.[8] Therefore, the Order applies to persons with mental health needs (in the language of the Bamford Review) or psycho-social disabilities, and people with learning, or intellectual disabilities. The Order specifically excludes the following individuals from its scope: those with 'personality disorder, promiscuity or other immoral conduct, sexual deviancy or dependence on alcohol or drugs'.[9]

The Order goes on to set out criteria for assessment (as to whether or not an individual should be detained and treated), involuntary detention and treatment in hospital, discharge, and procedural safeguards for detention and treatment – including recourse to the Mental Health Review Tribunal. In brief, an application for assessment can be made by an approved social worker or the person's nearest relative, and detention for the purposes of treatment can be authorised by a medical practitioner who certifies that the person has a mental illness or severe mental impairment which warrants detention, that failure to detain the patient would create a substantial likelihood of serious physical harm to himself or to other persons, and that no less restrictive alternatives to detention are available.[10]

The Order also provides for guardianship for persons with mental disorders, which is framed as a protective measure which is less restrictive than involuntary hospitalisation or imprisonment. Two types of guardianship are permitted under the Order – guardianship in s II provides for a patient over 16 (a person either with a mental disorder or 'appearing to be suffering from mental disorder') to be received into guardianship, where the disorder is of 'a nature or degree which warrants his reception into guardianship' and 'it is necessary in the interests of the welfare of the patient that he should be so received'. Under s III, courts can make a Guardianship Order in respect of persons with mental disorder who are convicted in criminal proceedings, as an alternative to imprisonment, and the guardian in these cases has the same powers as a guardian under s II.

Guardians have powers to require a person to reside at a particular place, to attend particular places at certain times for the purpose of medical treatment, occupation, education or training, and to provide access to the patient at any place where the patient is residing to any medical practitioner, approved social worker or other person.[11] These powers seem to be far-ranging, and there is little oversight of how they are exercised, and virtually no ability on behalf of individuals subject to guardianship to challenge the appointment of a guardian, or contest the need for guardianship in a court process. Guardians are initially appointed under the order for a period of six months, and orders can be renewed for one-year periods thereafter.[12]

The Code of Practice for the Mental Health Order describes the purpose of guardianship as follows: 'to ensure the welfare (rather than the medical treatment) of a

8 S 3(1), Mental Health (Northern Ireland) Order 1986.

9 Ibid s 3(2).

10 Ibid s 12(1).

11 Ibid s 22(1)(a).

12 Ibid ss 22(3) and 23(1).

patient in a community setting where this cannot be achieved without the use of some or all of the powers vested by guardianship'.[13] The Code also states that 'arrangements for giving effect to guardianship should not be unnecessarily complicated'[14] and that guardianship should be used 'in a positive and flexible manner'.[15] Further, it states 'if none of the powers given by guardianship are considered necessary for achieving the patient's welfare, then guardianship is inappropriate'.[16] Guardians are also required by the Code to be willing to 'advocate' for the patient to services needed to carry out the person's care plan.[17]

In relation to the powers of guardians, the Code makes clear that, although the powers specified in the Order seem far-ranging, a guardian is not empowered to use force or compel an individual under guardianship to act against their will. For example, it states that:

> the patient may be taken to the specified place if he willingly complies or offers no resistance. However, this power does not provide the legal authority to detain a patient physically in such a place, not does it authorize the removal of a patient against his will. If a patient is absent without leave from the specified place, he may be returned to it within 28 days by those authorized to do so[18]

– namely, the police, or an approved social worker. Finally, the Code states that 'if a patiently consistently resists the exercise of the guardian's powers, it can be concluded that guardianship is not the most appropriate form of care for that person and guardianship should be discharged'.[19]

It is clear that guardianship under the Mental Health Order allows for a form of substitute decision-making, albeit one that is intended, according to the Code, to be used sparingly. Although guardians appear to have powers to make decisions for those under their authority, guardians themselves cannot enforce these decisions. Although law enforcement and approved social workers can return an individual to a place specified by the guardian, there is no provision for guardians or others to override the individual's refusal of medical treatment, or to use force to gain access for professionals to an individual at their place of residence.

Other forms of substitute decision-making beyond guardianship in Northern Ireland are authorised under the common law – based on the judgments in *Re F*[20] and *Re C*.[21] The first judgment came in *Re F* in 1990, which stated that doctors have the power, and in certain circumstances the duty, to treat patients who lack mental capacity in their best interests. A clear definition of mental capacity was not provided in the case. Subsequently, in *Re C*, a person with schizophrenia refused an amputation of his gangrenous leg and the test of mental capacity the court applied was whether the individual could comprehend, take in and retain information, believe it, and weigh it up in order to make a decision. The court in *Re C* also emphasised that to have mental capacity an individual does not have to blindly accept medical evaluation and can have a level of self-assessment of any consequences. This definition is generally described as a functional assessment of mental capacity, and has subsequently been adapted and codified in the Mental Capacity Act 2005,

13 Department of Health and Social Services, *Mental Health (Northern Ireland) Order 1986: Code of Practice* (HMSO 1992) 27.

14 Ibid.

15 Ibid.

16 Ibid.

17 Ibid 28.

18 Ibid 33.

19 Ibid 24.

20 *Re F v West Berkshire Health Authority* [1990] 2 AC 1.

21 *Re C* [1994] All ER 118.

but as this only applies to England and Wales, the need for reform in Northern Ireland had to be addressed separately.

Beginning the reform process: the Bamford Review of Mental Health and Learning Disability

The first concrete discussions about what such reforms should look like were commenced by the Bamford Review of Mental Health and Learning Disability, which began its work in 2002. The Bamford Review was commissioned in the wake of the Northern Ireland Act 1998 and the Human Rights Act 1998, which set the parameters for reform of the legal framework affecting persons with mental health and learning disabilities. The scope of the review was broad-ranging, as it examined the need to ensure equality and social inclusion for people with mental health and learning disabilities, as well as proposals for delivering better services. The review lasted five years and produced a number of reports on specific disabilities and services (e.g. Alzheimer's, autism-spectrum disorders, learning disability and mental health, forensic mental health services, alcohol and substance misuse, etc.) as well as some general overviews of equality legislation and a framework for mental health service delivery.

The two key reports for the purpose of this discussion are the 2006 report on *Human Rights and Equality of Opportunity* and the 2007 report on *A Comprehensive Legislative Framework*. The 2006 report contained a number of recommendations on mental capacity and human rights, stating that substituted decision-making could be contrary to human rights in certain circumstances, e.g. 'in relation to a person who is acknowledged to have capacity; in relation to a person who is deemed incapable, but who actually has capacity; and inappropriate substitute decision-making in relation to a person who does not have capacity'.[22] Therefore, the 2006 report does not recognise that legal capacity is a universal right, not dependent on mental capacity, and similarly, does not acknowledge that substitute decision-making could de facto be considered a human rights violation. This follows the general trajectory of European Court of Human Rights jurisprudence on deprivation of legal capacity – which outlines concerns about the lack of procedural safeguards for individuals whose legal capacity is denied and is willing to criticise decisions made by substitute decision-makers which have led to bad outcomes for the individual (e.g. institutionalization in a restrictive and inappropriate setting), but has, to date, stopped short of finding that the denial of legal capacity in itself can constitute a violation of European Convention on Human Rights (ECHR) rights.[23]

Taking this as a starting point, the 2007 report goes on to set out key proposals for new legislation 'to provide equally for all circumstances in which an individual's autonomy might be compromised on health grounds'.[24] In short, this seeks to apply the same legal framework to all persons perceived to lack mental capacity, whether the source of this lack of mental capacity is from a learning disability, mental health condition, or other cognitive disability, and whether the decision in question relates to an individual's medical treatment, personal or welfare decisions, or financial decision-making. The report recommends the adoption of a 'fusion' approach to legislative reform in this field.

22 Bamford Review of Mental Health and Learning Disability (Northern Ireland), *Human Rights and Equal Opportunity* (Belfast 2006) 25.

23 See, for example, *Shtukaturov v Russia* (App No 44009/05) [2008] ECHR 223; *Stanev v Bulgaria* (App No 36760/06) [2012] ECHR 46; *DD v Lithuania* (App No 13469/06) [2012] ECHR 254; *X and Y v Croatia* (App No 5193/09) [2011] ECHR 1835; *Sykora v Czech Republic* (App No 23419/07) [2012] ECHR 1960; *Mihailovs v Latvia* (App No 35939/10) [2013] ECHR 65; and *Lashin v Russia* (App No 33117/02) [2012] ECHR 63.

24 Bamford Review of Mental Health and Learning Disability (Northern Ireland), *A Comprehensive Legislative Framework* (Belfast 2007) 3.

Essentially, this means that mental capacity regimes should not operate separately from mental health law which allows for involuntary detention and treatment on the basis of a 'mental disorder' and some criteria of risk (e.g. of harm to self and others). The idea is that, if the concept of mental capacity is taken seriously, and applied in the same way regardless of the individual's particular disability, then persons who are deemed to possess mental capacity (on the basis of a functional assessment of capacity) should be allowed to validly refuse medical treatment – including treatment for a mental health condition, even if there are risks involved. The result should be that involuntary treatment cannot be administered to a person who is found to have mental capacity to make a decision about consent to treatment, and that such treatment will only be given to persons who are found to lack mental capacity.

The fusion approach to mental health and mental capacity law adopted in the 2007 report is based in part on the work of Szmukler and Dawson,[25] who published an article on this subject in 2006, and seem to have significantly influenced the findings of the Bamford Review. Szmukler and Dawson begin from the premise that both physical and mental illnesses should be treated consensually, and that any non-consensual treatment (of either a physical or mental health condition) is only justified if the individual lacks mental capacity. Therefore, they propose abandoning the criteria of risk to self or others – primarily used to justify involuntary treatment in the context of mental health. This was heralded by many,[26] including the authors of the Bamford Review, as a progressive development which would respect the human rights of persons with mental health conditions on an equal basis with others, as it was thought to provide a less discriminatory basis for intervention than the mere presence of 'mental disorder' coupled with some level of risk.

However, this 'fusion' approach and the subsequent proposals for legislative reform, including those set out in the 2007 Bamford Report, were made in isolation from, and without full regard to, the emerging developments in international human rights law – particularly the UN Convention on the Rights of Persons with Disabilities, as will be discussed further below. Therefore, scholars such as Bartlett[27] have suggested that this kind of reform of 'mental capacity' should be approached with caution, as it may not yield the progressive, non-discriminatory results it promises. One of the problems identified by Bartlett is that the requirement in the assessment of functional capacity to understand and appreciate the nature and consequences of the decision is too flexible and 'has led to allegations that capacity means agreeing with the psychiatrist'.[28] Further, he suggests that:

> it is difficult to see that the incapacity can sensibly be separated from the mental disability, given that it is the mental disability that is the direct cause of the incapacity. Insofar as the use of the disability as a criterion is discriminatory, therefore, the use of incapacity as a detention criterion merely moves direct discrimination into indirect discrimination, and this is not really a significant advance.[29]

25 J Dawson and G Szmukler, 'Fusion of Mental Health and Incapacity Legislation' (2006) 188 British Journal of Psychiatry 504–09.

26 See, for example, N Rees, 'The Fusion Proposal: A Next Step' (paper delivered at the Rethinking Rights-Based Mental Health Laws Workshop, Monash University, Prato Campus, Italy, 22 May 2009); A Holland, 'The Model Law of Szmukler, Dawson and Daw: The Next Stage of a Long Campaign?' (2010) 20 Journal of Mental Health Law 63.

27 See P Bartlett, 'A Mental Disorder of a Kind or Degree Warranting Confinement: Examining Justifications for Psychiatric Detention' (2012) 16(6) International Journal of Human Rights 831–44.

28 Ibid 840.

29 Ibid 841.

While the Bamford Review proposed that a fusion approach be adopted, it also set down some clear criteria for respecting the human rights of people with mental health and learning disabilities in the context of decision-making capacity. The 2007 report made explicit that legislative reform should be based on four key principles as follows:

> Autonomy – respecting the person's capacity to decide and act on his own and his right not to be subject to restraint by others.
>
> Justice – applying the law fairly and equally.
>
> Benefit – promoting the health, welfare and safety of the person, while having regard to the safety of others.
>
> Least Harm – acting in a way that minimises the likelihood of harm to the person.

Arguably, these principles could still provide a sound basis for legislative reform which more closely complies with the emerging human rights perspective on universal legal capacity from the CRPD. However, this would require a significant shift in focus from the comprehensive framework set out in the Bamford Review. Instead of assessing an individual's capacity to make a particular decision, using the functional approach, and providing for substitute decision-making where an individual is found to lack decision-making capacity, a CRPD-compliant approach would entail using the four principles to ground a support model of legal capacity, and ultimately abolishing involuntary treatment in mental health as well as other types of non-consensual medical treatment.

In response to the publication of the Bamford Review, the Northern Ireland Executive produced a consultation plan, *Delivering the Bamford Vision*,[30] in 2008, and a follow-up action plan in 2009 based on the results of the consultation process.[31] In the original consultation paper, the Executive proposed introducing separate but parallel legislation on mental health and mental capacity respectively and this position was reflected in the policy consultation document on a legislative framework for a Mental Health and Mental Capacity Bill published by the Department of Health in 2009.[32]

However, following consultation on *Delivering the Bamford Vision* and the Department of Health's consultation document on the proposed legislative framework, the Executive revised this position in a subsequent action plan, which stated that the Department of Health would lead the way in drafting a comprehensive legislative framework to address mental health and mental capacity in a single Bill.[33] While the 2008 consultation document references work ongoing in the UK as a whole to ratify the CRPD – it does not connect obligations in the CRPD, particularly in Article 12, to the legislation being proposed, and the 2009 action plan did not reference the Convention at all.

Although the 2009 consultation paper on a legislative framework is clearly out of date since the agreement to publish a single Bill was reached, it does set out what a legislative framework would look like in Northern Ireland, on which more recent documents, including the Equality Impact Assessment of the proposed Bill, were based. For this reason, it is worth considering its key proposals here. With respect to capacity legislation, this

30 Department of Health, Social Services and Public Safety, *Delivering the Bamford Vision: The Response of the Northern Ireland Executive to the Bamford Review of Mental Health and Learning Disability – Consultation Document* (Belfast 2008).

31 Department of Health, Social Services and Public Safety, *Delivering the Bamford Vision: The Response of the Northern Ireland Executive to the Bamford Review of Mental Health and Learning Disability – Action Plan 2009–2011* (Belfast 2009).

32 Department of Health, Social Services and Public Safety, *Legislative Framework for Mental Capacity and Mental Health Legislation in Northern Ireland – A Policy Consultation Document* (Belfast 2009).

33 Ibid 27.

consultation document accepted that the four Bamford principles should be at the core of the legislative framework. While the consultation paper set out the need for mental capacity to be assessed, it does not provide much detail on who will assess capacity (except to say that where the decision involves consent to treatment, it should be assessed by a skilled professional). Similarly, although the document acknowledges that, where the person lacks mental capacity to make a decision, substitute decision-makers will come into play, no detail on who can be a substitute decision-maker (e.g. family member, professional) is provided.

The document also lists a number of powers and safeguards which will be included in the legislative framework, including: the opportunity to make advance decisions and lasting powers of attorney (LPAs) (where an individual has mental capacity at the time these are made); the appointment of deputies to make financial decisions on behalf of persons who lack mental capacity and do not have an advance decision or LPA; the creation of a new Office of Public Guardian; the statutory right for carers to have their views taken into consideration in substitute decision-making; advocacy services for those who lack capacity; a new offence of neglect or ill-treatment of persons who lack capacity; and deprivation of liberty safeguards for those who are placed in an institution without their consent and lack capacity to provide such consent.[34]

With regard to the mental health legislation, the consultation document proposed that assessment and treatment could continue to be carried out on a non-consensual basis if the person has a mental disorder, poses a significant threat to the health welfare and safety of themselves or others, and has impaired decision-making capacity in relation to treatment. Again, the failure to define impaired decision-making capacity leaves much uncertainty – particularly as is it unclear whether a person must be actually found to lack capacity to make the decision as to assessment and treatment, or if any deficit at all in the individual's decision-making skills could lead to involuntary assessment and treatment.[35]

Towards a single Bill on mental health and mental capacity in Northern Ireland

The Equality Impact Assessment of the Mental Capacity (Health, Welfare and Finance) Bill published in 2010 provides more clarity and detail on the provisions of the proposed legislation, and does so based on the concept of a single Bill. The assessment of mental capacity was clarified as being a two-stage functional test.[36] First, the individual had to have an impairment of, or disturbance in, the functioning of the mind or brain. Secondly, if, as a result of the impairment, the individual is unable to understand, retain, use and weigh information relating to the decision and communicate that decision to others, then the person will be found to lack mental capacity.[37] This two-step test is an exact replica of the assessment of capacity provided for in the Mental Capacity Act 2005 England and Wales.[38]

The Equality Impact Assessment also proposed to apply the new legislative framework to persons aged 16 and over,[39] although some exceptions were envisaged, e.g. in relation to contractual capacity (which is only recognised at 18) and the ability to establish an LPA (again, this would only be permitted at 18).[40] It also proposed that, while, as envisaged in Bamford, the provisions of the new legislative framework should be equally applicable to

34 Department of Health, Social Services and Public Safety (n 32) 9–10.

35 Ibid 13–14.

36 Department of Health, Social Services and Public Safety, *Mental Capacity (Health, Welfare and Finance) Bill Equality Impact Assessment* (Belfast 2010).

37 Ibid 5.

38 Ss 2 and 3, Mental Capacity Act 2005.

39 Department of Health, Social Services and Public Safety (n 36) 5.

40 Ibid 7.

persons in the criminal justice system, exceptions might be made for some individuals with a mental disorder in the criminal justice system who might possess mental capacity to make a decision regarding their medical treatment, but might be involuntarily treated nonetheless if they posed a substantial risk of serious harm to self and others.[41]

With regard to substitute decision-making where an individual is deemed to lack mental capacity, the Equality Impact Assessment proposed a four-pronged approach, based on the significance of the intervention to be made. The first, most common type of substitute decision is referred to as a 'routine' intervention and relates to the 'care, treatment and related expenditure which seek to meet the basic life and care needs of the person'.[42] A formal assessment of mental capacity will not be conducted as a basis for this kind of intervention – rather, the substitute decision-maker (e.g. carer or health professional) may act in the best interests of the individual if she has a reasonable, objective, belief that the person lacks mental capacity to make the decision in question. Any action taken must also be the least restrictive of the person's freedom, and the individual must be given the opportunity to participate in the decision, with the substitute decision-maker having due regard to 'their past and present wishes, beliefs and values and regard for the views of others, including carers'.[43] No automatic requirement to record substitute decisions made for routine interventions is envisaged in the Equality Impact Assessment.

The second category of substitute decision-making proposed relates to 'urgent interventions'. These are similar to routine interventions, but refer to situations where a decision needs to be made quickly in order to provide life-saving treatment. Due to the urgency of the situation, the Equality Impact Assessment states that 'it will not be appropriate for the person undertaking the intervention to undertake a full assessment of mental capacity'.[44] Therefore the substitute decision-maker is similarly empowered to act in the best interests of the individual if she has a reasonable belief that the individual lacks mental capacity. Substitute decision-makers must also take reasonable steps to establish if the individual has a lasting power of attorney and/or advance statement and take account of these in the decision to be made.

The third category of substitute decision-making set out in the Equality Impact Assessment refers to formal interventions in care, treatment, welfare or financial affairs 'which have significant consequences or are intrusive or restrictive'.[45] Due to the nature of these interventions, a formal assessment of mental capacity is proposed, based on the two-step test described above. The same safeguards apply as for less restrictive interventions and, in addition, a nominated person will be available to the person to ascertain her past and present wishes and ensure that these are communicated to the substitute decision-maker. If a nominated person is not identified or available, the individual is entitled to an advocate, who will fulfil the same role.

The final type of substitute decision-making relates to 'specifically authorised interventions' for care and treatment which have 'irreversible consequences, a high level of intrusiveness, complexity or risk, or are of a lengthy duration'.[46] These could include issues of organ donation, sterilisation, compulsory treatment for a 'mental disorder', deprivation of liberty, etc. In addition, the Equality Impact Assessment states that '[i]nterventions which

41 Department of Health, Social Services and Public Safety (n 36) 9.

42 Ibid 11.

43 Ibid.

44 Ibid 12.

45 Ibid.

46 Department of Health, Social Services and Public Safety (n 36) 12–13.

would normally fall under formal interventions but where there is resistance, objection or dispute, will also be dealt with under this level'.[47] Similar to formal interventions, these will require a formal assessment of mental capacity and written evidence supporting the intervention. The same safeguards as for formal interventions will be available to the individual, including a nominated person and/or advocate. The Equality Impact Assessment suggests that specific authorization for these interventions could be provided by a structure based on the existing Health and Social Care Trusts in Northern Ireland, overseen by the Regulation and Quality Improvement Authority. Authorisation for particularly serious interventions (involving life-saving treatment) may need to be provided for separately by the High Court.

There are a number of concerns with the proposed framework from a human rights perspective. Firstly, almost anyone the person interacts with can act as a substitute decision-maker, with very little oversight and procedural safeguards for the individual, including limited rights of appeal. No explicit mention is made in the Equality Impact Assessment of how a person subject to best-interests decision-making can challenge the validity of a best-interests substitute decision, or whether the intervention has been correctly classified (as routine, urgent, formal, or specifically authorised).

If, as it appears in the proposal, the individual's only recourse is to the High Court, this will make redress inaccessible for most people affected by substitute decision-making. Similarly, no reference to the right to instruct a solicitor appears in the proposals and so it is unclear whether capacity to instruct will be assessed, and by whom, or if solicitors can refuse the instructions of a client whom they believe lacks mental capacity. Northern Ireland has an Official Solicitor who 'provides legal representation for persons under a disability; this includes both minors (children under 18 years) and adults suffering from a mental disability (known as patients)'.[48] The ability to challenge a decision of the Official Solicitor (who can, for example, consent to treatment on behalf of 'patients') is not addressed in the Equality Impact Assessment – neither does it provide guidance on when the Official Solicitor should be appointed to individuals under the proposed new legislative framework.

The procedural, due process safeguards for individuals under substitute decision-making and for those who are deprived of their liberty by an intervention under the new legislative framework are also problematic. While the Equality Impact Assessment references the need to provide safeguards for those deprived of their liberty who may not have consented to placement in a residential setting (based on *HL v UK*),[49] it does not provide further detail on how such safeguards will be secured and, therefore, it appears that such deprivations of liberty would simply be authorised by a best-interests substitute decision. Specific deprivation of liberty safeguards do exist in the Mental Capacity Act 2005 (England and Wales) on which much of the Northern Ireland legislative framework is based – but these have been criticised for their piecemeal application[50] and concerns about effective monitoring of the safeguards have been repeatedly raised.[51]

47 Department of Health, Social Services and Public Safety (n 36) 13.

48 Northern Ireland Courts and Tribunals Service, *Royal Courts of Justice Customer Information: Official Solicitors Office* <www.courtsni.gov.uk/en-gb/aboutus/rcj/pages/royal%20courts%20of%20justice%20customer%20information.aspx#Solicitors> accessed 7 May 2013.

49 App No 45508/99 [2004] ECHR 471.

50 See R Hargreaves, *Deprivation of Liberty Safeguards: An Initial Review of Implementation* (Mental Health Alliance 2010).

51 See L Series, *Deprivation of Liberty Safeguards: A Haphazard Affair*, Social Care Network, Guardian Professional <www.guardian.co.uk/social-care-network/2012/apr/02/deprivation-liberty-safeguards-improvements?CMP=twt_gu> accessed 7 May 2013.

In short, the Equality Impact Assessment entirely envisages the establishment of a substitute decision-making system, based on a medical model of disability and a functional assessment of mental capacity. The supports envisaged in the Bill to support individuals in the exercise of legal capacity – advance statements, LPAs, the availability of advocacy and the role of the nominated person – are relatively weak. For example, advance statements are not legally binding, and must only be 'had regard to' not followed, even where the individual is deemed to possess mental capacity, and LPAs only take effect once an individual has 'lost' mental capacity and cannot be revoked or amended once they have come into effect. The supports envisaged in the legislative framework exist in parallel with substitute decision-making and do not intend to replace it. Finally, while the proposed legislative framework purports to abolish involuntary detention and treatment permitted under the current Mental Health Order 1986 it does not in fact achieve this, as it envisages that persons with mental disorders in the criminal justice system and children may still be detained and treated against their will, regardless of any assessment which demonstrates that the individual has mental capacity for the decision in question.

Consultation on legislative proposals and key concerns emerging

In December 2010, the Department of Health subsequently published a synthesis report of the submissions it received during the consultation process on the Equality Impact Assessment. In this document, the department reaffirmed that the single Bill approach would be based on the four Bamford principles of autonomy, justice, benefit and least harm, stating that 'the Autonomy principle is the foundation on which the legislation will be based'.[52] The document concluded that there was broad support for the department's proposal to legislate for a single Bill in this area, notwithstanding the complexity of the issues involved. Further detail was requested by many respondents in a number of areas including compulsory mental health treatment in the community, the application of the legislation in the criminal justice system, and the inclusion of persons with learning disabilities with 'abnormally aggressive or seriously irresponsible conduct' within the definition of mental disorder.[53]

One of the key findings of this document was the concern expressed by some respondents that, in terms of the legislative principles, there appeared to be a drift away from the 'Benefit' principle to one which emphasised 'Best Interests', which most felt was too paternalist in its focus.[54] Other key recommendations from respondents included the need for independent advocacy to be provided for all levels of intervention under the Act[55] (not just the more restrictive or intensive interventions) and the need for more due process protections and the ability to challenge both assessments of mental capacity and the decisions made under the various levels of intervention provided.[56]

Another key issue raised by submissions to the consultation paper was the treatment of children under the proposed legislative framework. Concerns were raised about the general exclusion of children under 16 from the Department's proposals, the failure to allow young people aged 16 and 17 to make lasting powers of attorney, and the exclusion of 16 and 17-

52 Department of Health, Social Services and Public Safety, *Mental Capacity (Health, Welfare and Finance) Bill Equality Impact Assessment – Analysis of Responses* (Belfast 2010) 3.

53 Ibid 6.

54 Ibid 8.

55 Ibid 10.

56 Ibid 9.

year-olds from the deprivation of liberty safeguards to be adopted in the new legislation.[57] Some submissions also requested a different kind of presumption of capacity for children – possibly adopting the 'mature minor' test as set out in *Gillick v West Norfolk*.[58] Consequently, it was recommended that the Children (NI) Order 1995 would have to be reviewed to bring it in line with the new legislative framework.[59]

The most recent consultation on the legislative proposals centred on the issue of extending the scope of the single Bill to the criminal justice system. A report on the responses to the consultation was published by the Department of Justice in January 2013. This report stated that many aspects of the criminal justice system in Northern Ireland already adopted a mental capacity approach – taking issues of decision-making capacity into account in determining the appropriate course of action, and welcomed the codification of this approach in the new Bill. The following issues were some of the key questions on which consultations were being sought:

> The issues which then arose for the justice system ranged across *day-to-day delivery* of services within the justice system – how a police officer for example might apply the capacity approach in his/her daily routines; what constitutes *best interests* for the individual where a crime has been committed and a judge is choosing a sentence; or where a person is in a *stressful custodial environment*, how will that influence decisions? *Protection* of the individual and others is a heightened factor in the justice system. What role might 'third parties' play in a system well versed with formal legal representation?[60]

In general, the findings of this consultation do not significantly depart from previous reports on the scope of the Bill – with the recommendation that most of the existing powers under the Mental Health Order in relation to criminal justice provisions should be retained.[61] While the report acknowledged that the courts would maintain their independence in terms of sentencing, it recommended that the defendant's mental capacity be taken into account in this process.[62] The report also recommended the retention of the police powers to remove a person suspected of suffering from a mental disorder to a place of safety, but also suggested that the places to which a person could be removed should be expanded beyond police stations, and that more community-based options should be available, including greater diversion opportunities for defendants with mental health problems.[63]

Critical analysis of the proposed legislation from a human rights standpoint: the CRPD

To date, there has been relatively little discussion about the need for the proposed legislative framework in Northern Ireland to comply with international human rights norms on legal capacity established in the CRPD. This is despite the ratification of the Convention and its Optional Protocol by the UK in 2009. The UK did not ultimately enter a declaration or reservation in relation to Article 12 on legal capacity, although it did consider a reservation

57 Department of Health (n 52) 6.

58 *Gillick v West Norfolk & Wisbeck Area Health Authority* [1986] AC 112.

59 Department of Health, Social Services and Public Safety (n 52) 7.

60 Department of Justice, *Consultation on Proposals to Extend Mental Capacity Legislation to the Criminal Justice System in Northern Ireland* (Belfast 2013) 10 (emphasis in original).

61 Ibid 3.

62 Ibid 4.

63 Ibid.

on Article 12(4)[64] and did enter declarations and reservations in respect of other issues, such as employment and education.

For clarity, it is worth setting out the key elements of Article 12 here, before examining the commentary to date on Northern Ireland's compliance with these provisions of the Convention. Article 12 states as follows:

> 1 States Parties reaffirm that persons with disabilities have the right to recognition everywhere as persons before the law.
>
> 2 States Parties shall recognize that persons with disabilities enjoy legal capacity on an equal basis with others in all aspects of life.
>
> 3 States Parties shall take appropriate measures to provide access by persons with disabilities to the support they may require in exercising their legal capacity.
>
> 4 States Parties shall ensure that all measures that relate to the exercise of legal capacity provide for appropriate and effective safeguards to prevent abuse in accordance with international human rights law. Such safeguards shall ensure that measures relating to the exercise of legal capacity respect the rights, will and preferences of the person, are free of conflict of interest and undue influence, are proportional and tailored to the person's circumstances, apply for the shortest time possible and are subject to regular review by a competent, independent and impartial authority or judicial body. The safeguards shall be proportional to the degree to which such measures affect the person's rights and interests.
>
> 5. Subject to the provisions of this article, States Parties shall take all appropriate and effective measures to ensure the equal right of persons with disabilities to own or inherit property, to control their own financial affairs and to have equal access to bank loans, mortgages and other forms of financial credit, and shall ensure that persons with disabilities are not arbitrarily deprived of their property.

The first paragraph reaffirms the status of persons with disabilities as persons before the law, i.e. as individuals possessing legal personality, with legal status and agency. This is confirmed by para 2, which extends the right to enjoy legal capacity on an equal basis with others to all aspects of life. Some might argue that functional assessments of mental capacity (which result in the removal of an individual's legal capacity in respect of a particular decision) conform with Article 12(2), since all adults, regardless of whether or not they have a disability, could, in theory, be subject to a functional assessment of their mental capacity, and have their legal capacity removed for a particular decision if they fail to meet a certain standard of decision-making ability.

However, since most functional assessments of capacity, including the one proposed in Northern Ireland's new legislative framework, require as a first step that the individual have an 'impairment in the functioning of the mind or brain', this means that, in reality, the functional assessment of mental capacity is disproportionately applied to people with disabilities, especially those with cognitive disabilities including psycho-social disability, intellectual disability, and people with dementia. Therefore, as Arstein-Kerslake and I have argued elsewhere,[65] when Article 12 is read in conjunction with Article 5 (Equality and

64 In the early stages of the ratification process, the UK government did consider entering a reservation to Art 12(4), on the basis that no review process was in place for appointees who collected payments from the Department of Work and Pensions on behalf of persons with disabilities who 'lack physical or mental capacity'. This reservation was subsequently dropped and not included in the final reservations and declarations submitted to the UN. See Department of Work and Pensions, *Explanatory Memoradum on the Ratification of the UN Convention on the Rights of Persons with Disabilities* (London, 3 March 2009).

65 E Flynn and A Arstein-Kerslake, 'The Support Model of Legal Capacity: Fact, Fiction, or Fantasy?' (2013) 30(4) Berkeley Journal of International Law (forthcoming).

Non-Discrimination), the use of functional assessment of mental capacity to justify substitute decision-making is discriminatory (in both purpose and effect) towards persons with disabilities. This interpretation is now reflected in the Committee's Draft General Comment on Article 12.[66]

Paragraph 3 contains one of the more novel additions to Article 12 – a state obligation to provide the supports required to exercise legal capacity. As discussed in the Introduction above, this support can take many forms, including, but not limited to, formal agreements with supporters who assist in certain areas of decision-making. Paragraph 4 addresses the safeguards required for all measures regarding the exercise of legal capacity. Some argue that this provision allows for some limited forms of guardianship to remain if the appropriate safeguards are in place;[67] however, the Committee on the Rights of Persons with Disabilities has not accepted this argument from any of the countries it has examined to date – even those countries which have interpretative declarations on Article 12, setting out interpretations that Article 12 permits some limited forms of substituted decision-making.[68]

The key phrase in Article 12(4), and one which has been used repeatedly by the Committee on the Rights of Persons with Disabilities in its concluding observations on the countries examined to date, is that safeguards should be designed to respect the 'rights, will and preferences' of the person. The term 'best interests' does not appear in para 4 , or in Article 12 at all. Therefore, I would argue that safeguards which are paternalistic in nature, or which envisage the use of substitute decision-making are not permitted under Article 12. Finally, para 5 refers specifically to the need to respect legal capacity with regard to financial affairs and property – an issue which was subject to extensive debates during the negotiation of the CRPD.

The Committee has repeatedly called for the abolition of regimes of substitute decision-making and their replacement with systems of supported decision-making in each of the seven concluding observations it has issued to date.[69] In each of its Concluding Observations on these countries, the Committee expressed concern 'that no measures have been undertaken to replace substitute decision-making by supported decision-making in the exercise of legal capacity'.[70] With respect to all countries, the Committee recommended that the states 'review the laws allowing for guardianship and trusteeship, and take action to develop laws and policies to replace regimes of substitute decision-making by supported decision-making, which respects the person's autonomy, will and preferences.'[71] This

66 Committee on the Rights of Persons with Disabilities (CRPD), *Draft General Comment on Article 12: Advance Unedited Version*, 10th Session (2–13 September 2013), paras 9, 21 and 28.

67 See Canada's Reservation to the CRPD, Art 12 (11 March 2010). However, the United Nations Committee on the Rights of Persons with Disabilities (CRPD) has recently noted in several of its concluding observations that steps should be taken to replace substituted decision-making with supported decision-making: Committee on the Rights of Persons with Disabilities (CRPD) *Tunisia* (n 4); Committee on the Rights of Persons with Disabilities (CRPD) *Spain* (n 4).

68 See, for example, *Consideration of Reports Submitted by States Parties under Article 35 of the Convention: Concluding Observations, Australia*, Committee on the Rights of Persons with Disabilities (CRPD), 10th Session, UN Doc CRPD/C/AUS/CO/1 (2–13 September 2013).

69 See, for example, Committee on the Rights of Persons with Disabilities (CRPD) *Tunisia* (n 4); Committee on the Rights of Persons with Disabilities (CRPD) *Spain* (n 4).

70 *Consideration of Reports Submitted by States Parties under Article 35 of the Convention: Concluding Observations of the Committee on the Rights of Persons with Disabilities*, Committee on the Rights of Persons with Disabilities, 5th session, 4 (11–15 April 2011) <www.ohchr.org/EN/HRBodies/CRPD/Pages/Session5.aspx> accessed 26 April 2013.

71 *Consideration of Reports Submitted by States Parties under Article 35 of the Convention: Concluding Observations of the Committee on the Rights of Persons with Disabilities*, Committee on the Rights of Persons with Disabilities, 6th session, 5 (19–23 September 2011) <www.ohchr.org/EN/HRBodies/CRPD/Pages/Session6.aspx> accessed 17 May 2013.

approach demonstrates the Committee's acceptance of the need for a support model of legal capacity to be implemented in States Parties to the Convention; and following the publication of the Committee's Draft General Comment more guidance has been provided on the definitions of 'substitute decision-making regimes' and 'supported decision-making' respectively. Substitute decision-making is defined as follows:

> a system where 1) legal capacity is removed from the individual, even if this is just in respect of a single decision, 2) a substituted decision-maker can be appointed by someone other than the individual, and 3) any decision made is bound by what is believed to be in the objective 'best interests' of the individual – as opposed to the individual's own will and preferences.[72]

By contrast, the Committee provides a broad interpretation of 'supported decision-making', as 'a cluster of various support options which give primacy to a person's will and preferences and respect human rights norms'.[73] A non-exhaustive list of support options is provided in the draft General Comment, from relatively minor accommodations, such as accessible information, to more formal measures such as supported decision-making agreements nominating one or more supporters to assist the individual in making and communicating certain decisions to others.[74] The Committee also clarifies that legal capacity is a universal attribute,[75] and that it cannot be made contingent on the individual's level of mental capacity.[76]

Further to the Committee's interpretation, and building on the work of Bach and Kerzner, Arstein-Kerslake and I have suggested some additional components of the support model of legal capacity as enshrined in Article 12. We argue that it is crucial that supports for exercising legal capacity could be offered to the individual, but not imposed.[77] In situations where the person's will and preferences remain unknown after significant efforts to discover these have been made, 'facilitated' decision-making[78] could be available, where someone could be appointed to make a decision on behalf of another individual as a last resort. In all situations where support is provided, the most important safeguard, as has been emphasised by the Committee, is the need for decisions to reflect the rights, will and preferences of the individual receiving support.[79] Having considered the possible implications and interpretations of Article 12, I will now turn to the UK position on legal capacity and its obligations under the CRPD.

In the UK's initial report to the Committee on the Rights of Persons with Disabilities, the section on Article 12 states that: 'The UK Government strongly supports the equal recognition of disabled people before the law, and their right to exercise legal capacity.'[80] It is important to note also that the UK government position on treaty ratification is that it will only ratify once all necessary laws have been enacted to ensure conformity with the

72 Committee on the Rights of Persons with Disabilities (CRPD), *Draft General Comment on Article 12* (n 66) para 23.

73 Ibid para 25.

74 Ibid para 15.

75 Ibid para 8.

76 Ibid para 12.

77 Flynn and Arstein-Kerslake (n 65) 5.

78 The concept of facilitated decision-making was conceptualized by Michael Bach and Lana Kerzner in 'A New Paradigm for Protecting Autonomy and the Right to Legal Capacity' prepared for the Law Commission of Ontario (October 2010) <www.lco-cdo.org/disabilities/bach-kerzner.pdf> accessed 17 May 2013.

79 Committee on the Rights of Persons with Disabilities (CRPD), *Draft General Comment on Article 12* (n 66) para 15.

80 Office for Disability Issues, *UK Initial Report on the UN Convention on the Rights of Persons with Disabilities* (London 2011) 17.

Convention.[81] However, the report also states that: 'There may be circumstances in which a disabled person needs support to exercise that capacity, or *where they lack the mental capacity to make decisions for themselves, and decisions may then be made on their behalf*.'[82] The highlighted portion of the second statement clearly indicates that the UK, including Northern Ireland, has taken the stance that substitute decision-making is permissible for persons with disabilities who are deemed to 'lack mental capacity' – usually according to the functional test as explained in the Introduction above.

This does not, however, accord with the Committee's own statements on substitute decision-making which will be discussed further below. There is only one reference to Northern Ireland in the Article 12 section of the UK report, which states that: 'In **Northern Ireland**, the Mental Health (NI) Order 1986 currently provides a framework for the Courts to manage and administer the finance, property and affairs of adults who lack capacity.'[83] No further discussion of the reforms related to Article 12 underway in Northern Ireland, including the proposal for a single Bill on mental health and mental capacity, and whether these reforms will comply with the CRPD, is made in the UK report.

The only shadow report published to date on the UK, written by the Deaf Ex-Mainstreamers Group,[84] focuses on Article 24 (Inclusive Education) and does not address the issue of legal capacity. An expert paper on disability policy and programmes in Northern Ireland was commissioned by the Equality Commission and published in 2011.[85] This paper covers all articles of the Convention, analysing the relevant policies and programmes in Northern Ireland and their compliance with the Convention, and combining this with insights from focus groups of disabled people. Although the section of this expert paper examining Article 12 is relatively brief, it does make some important points. With regard to Article 12(3), the paper states as follows:

> Whilst the requirements of Article 12(3) may be seen as resource intensive, there is no reason to believe that the right to access the support needed to exercise one's legal capacity is a socio-economic right and thus is a progressively realisable obligation on States Parties as per CRPD Article 4(2). Recognition as a person before the law is an example of a classic civil and political right and it is of immediate effect. With respect to persons with disabilities, the obligation to take measures to ensure access to support under Article 12(3) is also of immediate effect. This means that Article 12(3) is not an obligation to provide access only through taking 'measures to the maximum of its available resources' (Article 4(2)). Article 12(3) support must also not be seen as a form of 'reasonable accommodation' which is subject to limitation in 'not imposing a disproportionate or undue burden' (Article 2).[86]

The paper also supports the position that distinctions in the enjoyment of legal capacity which are based on disability are discriminatory and contrary to the spirit and intention of the CRPD. Based on my argument above regarding the discriminatory nature of functional assessments of capacity (especially where these tests are premised on the existence of an

81 Foreign and Commonwealth Office, *Treaties and MOUs: Guidance on Practice and Procedures* (London 2012) 6.

82 Office for Disability Issues, *UK Initial Report on the UN Convention on the Rights of Persons with Disabilities* (London 2011) 17, emphasis added.

83 Ibid 18 (bold in original).

84 Deaf Ex-Mainstreamers Group, *Shadow Report to the UN Secretary General and the Committee on the Rights of Persons with Disabilities* (London 2011).

85 C Harper, S McClenahan, B Byrne and H Russell, *Disability Programmes and Policies: How Does Northern Ireland Measure Up? Monitoring Implementation (Public Policy and Programmes) of the United Nations Convention on the Rights of Persons with Disabilities (UNCRPD) in Northern Ireland* (Equality Commission 2011).

86 Ibid 124.

impairment in functioning of the mind or brain), this would seem to indicate that the current legislative framework on mental capacity being proposed in Northern Ireland should be revised. Finally, the results of the questionnaire distributed by the authors of the expert paper indicated that mental health and capacity was the fourth highest priority for change in light of the CRPD for non-disabled respondents and the sixth highest priority for change overall.[87] The comments made by respondents on this issue focused on stigma, knowledge deficits and discrimination, but few identified what kind of reform they wished to see, which is perhaps unsurprising, given the complexity of this issue and the lack of campaigning from the disability movement or civil society as a whole in Northern Ireland on the need to move away from substitute decision-making.[88]

No statement regarding Northern Ireland's position on Article 12 and the proposed legislative reforms has been made by either the Equality Commission of Northern Ireland, or the Northern Ireland Human Rights Commission. However, the Independent Mechanism established under Article 33 of the CRPD to monitor the implementation of the Convention at domestic level (a task shared by the Human Rights and Equality Commissions in Northern Ireland) did provide advice to government on the scope of the proposed legislation in light of the CRPD, and its application to children in particular.[89] While this report does not challenge the central assumption that functional assessments of mental capacity which result in the removal of legal capacity are contrary to Article 12, it does refer to the need for the legislation to focus more squarely on 'supported decision-making' rather than simply providing for different kinds of 'substitute decision-making' to occur:

> There is a contrast between Article 12, which emphasises supported decision making and the Bill, which establishes a system of substitute decision making. Article 12 aims to facilitate decision making *with* the person, rather than for the person. In this context, the Bill must ensure that persons experiencing difficulties with decision making have access to the support necessary to make a decision themselves.[90]

This statement is one of the only examples of critique of the legislative proposals in Northern Ireland to date from the standpoint of Article 12. Some awareness-raising about the need to further reform the legislative proposals was conducted by Mencap, which organised a roundtable on legal capacity in conjunction with Inclusion Europe, Inclusion Ireland, the Law Centre of Northern Ireland and the Centre for Disability Law and Policy, National University of Ireland Galway, in November 2012.[91] While officials from the Departments of Justice and Health attended the roundtable, the impression they gave was that the legislation was almost finalised and, subject to the consultation discussed above about the extension of the proposals to the criminal justice system, the legislative framework would remain as set out in the Equality Impact Assessment in 2010, with a functional assessment of mental capacity, and four different types of interventions allowing for substitute decisions to be made in the best interests of a person deemed to lack sufficient mental capacity.

Therefore, it seems that the approach to be taken in Northern Ireland will follow closely the legislation currently in force in England and Wales (Mental Capacity Act 2005) and

87 Harper et al (n 85) 127.

88 Ibid.

89 UN Convention on the Rights of Persons with Disabilities Independent Mechanism for Northern Ireland, *Advice to Government on the Mental Capacity (Health, Welfare and Finance) Bill* (Belfast 2012).

90 Ibid 3.

91 See Mencap Northern Ireland, 'An Opportunity to Turn Rights into Reality' press release, 26 November 2012 <www.mencap.org.uk/news/article/opportunity-turn-rights-reality> accessed 27 May 2013.

Scotland (Adults with Incapacity (Scotland) Act 2000). All three are premised on a functional assessment of mental capacity which leads to the subsequent removal of legal capacity in respect of a particular decision, and substitute decision-making which follows a best-interests approach. As discussed above, such systems have already been described by other scholars as potentially conflicting with the requirements of Article 12 CRPD.[92] In addition, the paternalistic approach of the Court of Protection in England and Wales to the authorisation of best-interests decision-making and the lack of due process safeguards and access to justice for individuals subject to substitute decision-making in these systems has also been criticised.[93]

In the following section, I contrast this position with current developments in legal capacity legislation in the Republic of Ireland, with a view to advocating that a consistent approach be taken in both jurisdictions, in order to ensure equivalence of human rights protection on both sides of the border, in light of the Belfast/Good Friday Agreement.

Legal capacity reform in Ireland: implementing the CRPD

The Republic of Ireland has committed to reform its outdated substitute decision-making regime, known as the ward of court system, prior to ratifying the Convention. The current ward of court system in Ireland encapsulates a status-based approach to legal capacity, whereby a person of 'unsound mind' who is 'incapable of governing his person or property' is deprived of all legal rights in respect of decision-making.[94] A Law Reform Commission Report published in 2006[95] (prior to the adoption of the CRPD) recommended the introduction of a functional assessment of mental capacity, and the subsequent appointment of personal guardians if an individual was found to lack sufficient mental capacity for the relevant decision. However, since the entry into force of the Convention, discourse in the Republic about the possible scope and outlook of new capacity legislation has moved away from substitute decision-making towards a support model of legal capacity.

The present Minister for Justice, Alan Shatter, made the following statement in Parliament to the effect that Ireland would not ratify until the necessary legislative reforms were completed: 'Ireland does not become party to treaties until it is first in a position to comply with the obligations imposed by the treaty in question, including by amending domestic law as necessary.'[96] When the present government came to power in 2011, its Programme for Government contained a commitment to introduce a 'Capacity Bill that is in line with the UN Convention on the Rights of Persons with Disabilities'.[97]

The parliamentary joint committee on Justice, Defence and Equality ('the Justice Committee') subsequently called for submissions in August 2011 from interested parties on the content of what was being referred to at the time as the Mental Capacity Bill.[98] In response, a coalition of interested organisations and individuals, in the fields of intellectual

92 See P Fennell and U Khaliq, 'Conflicting or Complementary Obligations? The UN Disability Rights Convention, the European Convention on Human Rights and English Law' (2011) (6) European Human Rights Law Review 662–74; Bartlett (n 27).

93 See Series (n 51).

94 Lunacy Regulation Act 1871 (c 22/1871) (Ir).

95 Law Reform Commission, *Report: Vulnerable Adults and the Law* (Law Reform Commission 2006).

96 Dail Debates, 22 May 2012, Written Answers: National Disability Strategy <http://debates.oireachtas.ie/dail/2012/05/22/00318.asp> accessed 26 April 2013.

97 Department of An Taoiseach, *Programme for Government 2011* (Government Publications 2011) 8 <www.taoiseach.gov.ie/eng/Publications/Publications_Archive/Publications_2011/Programme_for_Government_2011.pdf> accessed 26 April 2013.

98 Oireachtas Joint Committee on Justice, Defence and Equality, *Mental Capacity Legislation: Invitation for Submissions* (August 2011) <www.oireachtas.ie/parliament/mcl/> accessed 12 February 2013.

disability, mental health, and older people, co-chaired by the Centre for Disability Law and Policy, National University of Ireland Galway, and Amnesty International Ireland's Mental Health Campaign, came together to discuss whether a joint approach to legal capacity reform could be developed across these interest groups.

The result was the publication of an agreed set of Essential Principles for Legal Capacity Reform in April 2012, setting out 10 key principles to be adhered to in legislation designed to comply with Article 12 CRPD.[99] Key requirements in these principles included the need for support to fully replace substitute decision-making, the abolition of best interests as a criteria in decision-making and its replacement with an approach that respected the 'will and preferences of the person' and the state obligation to provide a range of support options to individuals who wish to use support in exercising their legal capacity. Many of the groups involved presented at oral hearings convened by the Justice Committee in February 2012.[100] The Justice Committee also published a report based on the oral hearings, endorsing the support model of legal capacity and requiring a shift away from the best-interests model of substitute decision-making towards an approach which respects the will and preferences of the individual.[101]

> To begin with, what is needed is a starting point whereby we accept that all persons have legal capacity and are capable of making decisions regarding what to do in all areas of life . . . It is essential that the new legislation ensures that persons with disabilities have access to supports which allow them to fully enjoy all human rights. We as a society must move away from paternalism and fully recognise the 'will and preferences' of the individual as required by Article 12(4) of the UN Convention on the Rights of Persons with Disabilities.[102]

The Legal Capacity Bill was suggested as a more appropriate title for this legislation as required by international human rights law. Legal capacity – the capacity to hold and exercise rights – should be distinguished from mental capacity – which refers to decision-making capability. The Convention does not in any way mention 'mental capacity'.[103]

The Assisted Decision-Making (Capacity) Bill was published in July 2013. It presents an interesting mix of supports (including the option of entering binding assisted decision-making agreements[104] and co-decision-making agreements)[105] and substitute decision-making (such as decision-making representatives[106] and informal decision-makers),[107] but continues to be premised on the individual reaching a certain standard of mental capacity as a prerequisite for retaining legal capacity in respect of a given decision. The definition of capacity does not include a diagnostic step (i.e. impairment in the functioning of the mind or brain) which, on one view, makes it less obviously discriminatory, but on the other hand, any of the forms of decision-making prescribed under the Bill may only occur where the

99 Amnesty Ireland and Centre for Disability Law and Policy, *Essential Principles for Legal Capacity Reform* (2012) <htwww.nuigalway.ie/cdlp/documents/principles_web.pdf> accessed 26 April 2013.

100 Oireachtas, 'Committee on Justice, Defence and Equality Continues Hearings on Proposed Mental Capacity Legislation' press release (February 2012) <www.oireachtas.ie/parliament/mediazone/pressreleases/2012/name-6931-en.html> accessed 26 April 2013.

101 Oireachtas Joint Committee on Justice, Defence and Equality, *Report on Hearings in Relation to the Scheme of the Mental Capacity Bill* (May 2012) <www.oireachtas.ie/parliament/media/michelle/Mental-capacity-text-REPORT-300412.pdf> accessed 28 May 2013.

102 Ibid 7.

103 Ibid 8.

104 Assisted Decision-Making (Capacity) Bill 2013, s 10.

105 Ibid s 18.

106 Ibid s 24.

107 Ibid s 53.

individual considers that her capacity is either 'in question' or 'shortly may be in question'[108] which seems to imply that the main group of individuals affected by the legislation will be those with impaired decision-making ability – especially persons with cognitive disabilities.

A detailed discussion of the legislation is outside the scope of this paper, but it is important to note that, even in the substitute decision-making provisions of the Bill, intervenors are obliged to act in conformity with the guiding principles of the Bill, which include respect for the will and preferences of the individual (albeit with the qualifier that this should only be done 'as far as practicable').[109] It is also significant that best interests does not appear as a principle for guiding decision-making under the Bill.

The definition of capacity set out in s 3 of the Bill reveals that the underlying premise of the legislation is that a certain standard of mental capacity is a prerequisite for the recognition of an individual's legal capacity – a premise which is not, in my view, compatible with the Committee on the Rights of Persons with Disabilities' interpretation of Article 12. Nevertheless, despite this view, legal recognition of various supports necessary to exercise legal capacity (such as assisted decision-making and co-decision-making) are provided in the Bill, which is certainly to be welcomed.

At a minimum, there has been more awareness-raising and recognition in the Republic (than in Northern Ireland) of the need for legislation on capacity to abolish substitute decision-making in order to comply with the provisions of Article 12. In light of the commitment in the Belfast/Good Friday Agreement to ensure equivalence of protection for human rights on both sides of the border,[110] therefore, I suggest that the more progressive approach which appears to be the focus of new legislation in the Republic should be reflected in the development of legislation on capacity in Northern Ireland.

Conclusion

Since the entry into force of the CRPD, many countries are struggling to come to terms with the reforms necessary – and Article 12 on legal capacity is one of the most complex areas of law in need of reform. Members of the UN Committee on the Rights of Persons with Disabilities have repeatedly stated that no country is currently fully in compliance with Article 12, since the full replacement of substituted decision-making with supported decision-making has not yet been achieved.[111] Nevertheless, the Committee continues to require States Parties to take steps to eliminate regimes of substitute decision-making and replace these with supports to exercise legal capacity. Some important reform processes are currently underway throughout the world, including in the Republic of Ireland, as discussed above, and also in India[112] and the Canadian province of Newfoundland and Labrador.[113]

108 Assisted Decision-Making (Capacity) Bill 2013, s 2 (see definition of 'relevant person').

109 Ibid s 8.

110 See Cinnéide (n 5).

111 See, for example, T Degener, 'Monitoring the UN Convention on the Rights of Persons with Disabilities' (keynote address at Getting Ready to Ratify: The UN Convention on the Rights of Persons with Disabilities, Dublin, 17 May 2013).

112 See Committee appointed by Ministry of Social Justice and Empowerment, Government of India, *The Rights of Persons with Disabilities Bill 2011* (Centre for Disability Studies, NALSAR University of Law, Hyderabad 2011) <http://socialjustice.nic.in/pdf/report-pwd.pdf> and Draft National Trust Act Amendments 2011 <http://thenationaltrust.co.in/nt/images/national%20trust%20amendment.pdf> accessed 28 May 2013.

113 See Newfoundland and Labrador Association of Community Living, *Policy Document Submitted to Justice Aims to Aid People with Intellectual Disabilities* <www.nlacl.ca/news/article/getting-power-make-decisions-policy-document-submi/> accessed 26 April 2013.

It is therefore surprising that the requirements of Article 12 have had so little influence on the development of the Mental Capacity (Health, Welfare and Finance) Bill in Northern Ireland. With the exception of Mencap and some isolated disability organisations, there has been relatively little campaigning from civil society on the need for the new legislation to move away from substitute decision-making in order to comply with the CRPD. Similarly, the independent mechanism which oversees domestic implementation of the CRPD, appears to have only published one set of comments on the legislation which refer to the need for it to be CRPD-compliant – but does not challenge the central premise of the legislation, which is based on a functional assessment of mental capacity and the subsequent removal of legal capacity for a particular decision. While legislation in the Republic may not ultimately comply with all the requirements of Article 12 – at least there has been open acknowledgment from government on the need for the legislation to enable ratification of the CRPD[114] and repeated calls from civil society to place supported decision-making at the heart of the new Bill.[115]

In this paper I have sought to draw out the discontinuities between the legislative proposals in Northern Ireland, the requirements of Article 12 CRPD and the attempts to legislate for supported decision-making in the Republic of Ireland. The arguments presented will, I hope, serve to open further cross-border collaboration and discussion at both government and civil society levels on the need to uphold the fundamental rights of persons with disabilities – especially the right to legal capacity – one of the core human rights necessary for the enjoyment and exercise of almost all other rights.

110 Dail Debates (n 96).

111 Amnesty Ireland and Centre for Disability Law and Policy, *Essential Principles for Legal Capacity Reform* (2012) <www.nuigalway.ie/cdlp/documents/principles_web.pdf> accessed 26 April 2013.

NILQ 64(4): 507–23

Right answers and Realism: Ronald Dworkin's theory of integrity as a successor to Realism

STEPHEN W SMITH*

University of Birmingham

1 Introduction

To the extent that there is such a thing as American Jurisprudence,[1] Legal Realism[2] lies close to its foundation. Several critical schools of thought (e.g. Critical Legal Studies (CLS)) have claimed they are the progeny of Realism.[3] Other, more scientifically based schools, such as jurimetrics or the law and economics movement, are believed to derive from central claims of Legal Realism.[4] Other scholars, for example Brian Leiter, claim to continue to follow its tenets even if the version has been reinterpreted.[5] Legal Realism, then, accounts for the beginnings of a number of distinctly American legal theories. However, one of the most recognised American theories – Ronald Dworkin's law as integrity – does not make similar statements. Dworkin has never claimed Legal Realism as an influence although he is aware of its theories and has engaged with them.[6] Indeed, Dworkin and Legal Realism are rarely treated together at all.[7] To the extent that they are, they are often put on polar opposites specifically in relation to the indeterminacy of legal rules. In fact, Leiter has gone so far as to claim Dworkin's view is a 'sophisticated' view of Formalism.[8]

This article will contest that view. It will argue that Dworkin's theory of integrity follows on from the important tenets of Legal Realism. His theory of constructive interpretation,

* I would like to thank Gavin Byrne, Sean Coyle, James Lee, Sheelagh McGuinness, Frederick Schauer and Gordon Woodman for helpful comments on earlier drafts of this article. All the usual disclaimers apply.

1 N Duxbury, *Patterns of American Jurisprudence* (Clarendon Press 1995). For an opposing view, see B Leiter, 'Is there an "American" Jurisprudence' (1997) 17 OJLS 367–87 reprinted in B Leiter, *Naturalizing Jurisprudence* (Oxford University Press 2007) 81–102.

2 There is, of course, a Scandinavian Legal Realism in addition to the American version. However, the Scandinavian version is not one I will consider further. Thus, any reference to Legal Realism or Realism should be taken to mean exclusively American Legal Realism.

3 A Altman, 'Legal Realism, Critical Legal Studies, and Dworkin' (1986) 15 Philosophy and Public Affairs 205–35, 207.

4 Leiter (n 1) 95–6. But see, Duxbury (n 1) ch 5 for an alternative view.

5 Leiter (n 1) 21.

6 Pragmatism, as he calls it in *Law's Empire*, appears to be a form of Legal Realism as an interpretive theory: R Dworkin, *Law's Empire* (Belknap Press 1986) 151–64.

7 Frederick Schauer, however, has recently suggested that Dworkin and the Realists are likely to reach similar decisions in practice: F Schauer, 'Legal Realism Untamed' (2013) 91 Texas Law Review 749.

8 B Leiter, 'Legal Formalism and Legal Realism: What is the Issue?' (2010) 16 Legal Theory 111–33, 112.

while it builds upon and develops important Realist themes, requires a fundamental acceptance of some of its major conclusions. I wish to be clear what the claim entails. I am not suggesting that Ronald Dworkin is a Realist. He is not. Nevertheless, I will argue Dworkin and the Realists would find common ground in a number of important concepts. These concepts were important central tenets of Realism which Dworkin, explicitly or otherwise, has accepted and play foundational roles in his theory of law as integrity. There are crucial aspects where he deviates from Realism but Dworkin and the Realists have more in common than most think. It should also be made clear that this article will not present the claim that if one starts out with a Realist position, then one *must* end up with Dworkin's theory of integrity. It does not. The claim being made here is a more limited one. Dworkin's theory of law as integrity is one possible conclusion which can be reached from Realist premises. It is not the only possibility and nothing I say here should detract from the suggestions about other theorists like Leiter or those in the Critical Legal Studies movement that their theories are based upon Legal Realism. Realism is a broad church and while CLS and Leiter can rightly claim to be successors to the Realists, Dworkin could likewise claim inspiration from their views should he wish to do so.

To explore these connections it becomes necessary to determine what the central tenets of Realism are. This has never been an easy task as the Realists did not all derive their inspiration from one central source material nor did they all believe or argue the same sorts of things.[9] It covers everything from the strict behavioural account of legal practice used by Underhill Moore, to the sociological account utilised by Karl Llewellyn, to the Freudian psychology-based approach championed by Jerome Frank.[10] They did not always agree on everything and even major points of criticism of others applied as much to colleagues within the Realist camp as they did to those outside.[11] Despite this, there are some things which the Realists had in common and the first task of this article will be to explore those central ideas. This will be the focus of Part 2 of the article. Part 3 will then go on to explore some of the important theoretical questions which are raised by the Realists but which they never completely address. Part 4 will then examine how Dworkin's theory of integrity answers some of these resulting issues. The final section will examine what this means for our understandings of both Dworkin's theory as well as American Legal Realism.

2 Central tenets of American Legal Realism

As noted above, determining the central tenets of American Legal Realism is difficult. Due to the varied interests of those within the camp of American Legal Realism, it can be hard to credit anything as an actual central tenet. Underhill Moore, for example, seemed to spend a large majority of his time involved in describing behaviour which was some way related to legal rules.[12] Frank spent most of his time discussing either fact scepticism or the psychological implications of the law.[13] Others had similarly narrow interests even if some of them were more legal in origin.

Karl Llewellyn, however, did indicate that there were a number of characteristics of all Realists. He argued that they shared a number of things in common which he sets out in

9 Duxbury (n 1) 68.

10 Leiter (n 1) 16, 28–29.

11 Duxbury (n 1) 69; K Llewellyn, 'Some Realism about Realism: A Response to Dean Pound' (1930–1931) 44 Harvard Law Review 1222–64, 1233–34.

12 U Moore and C C Callahan, 'Law and Learning Theory: A Study in Legal Control' (1943) 53 Yale Law Journal 1–136.

13 J Frank, *Law and the Modern Mind* (Peter Smith 1970; original publication 1930); J Frank, *Courts on Trial* (Princeton University Press 1949).

'Some Realism about Realism' but only five of the nine suggested tenets are 'characteristic marks' of Realist scholars.[14] Those five are the following:

1 'the temporary divorce of "Is" and "Ought" for purposes of study';
2 . . . 'a distrust of the theory that traditional prescriptive rule-formulations are the heavily operative factor in producing court decisions';
3 'the belief in the worthwhileness of grouping cases and legal situations into narrower categories than has been the practice in the past';
4 'An insistence on evaluation of any part of law in terms of its effects, and an insistence on the worthwhileness of trying to find these effects'; and
5 'an insistence on sustained and programmatic attack on the problems of law along any of these lines'.[15]

Brian Leiter has likewise argued that there are central tenets of American Legal Realism. Leiter argues the 'Core Claim' of Realism is that 'judges respond primarily to the stimulus of facts'.[16] Leiter wants to draw particular attention to the word 'primarily'. He highlights that the Realists did not claim that rules were irrelevant to a judicial decision but that the main thrust of any judicial decision is what appears to be most fair under the factual circumstances of the case.[17] This Core Claim is further bolstered by what Leiter refers to as the Determination Thesis and the Generality Thesis.[18] The end result of this is that Realists, according to Leiter, are proponents of the following:

> (1) a descriptive theory about the nature of judicial decision, according to which, (2) judicial decisions fall into (sociologically) determined patterns, in which (3) judges reach results based on a (generally shared) response to the underlying facts of the case, which (4) they then rationalize after-the-fact with appropriate legal rules and reasons.[19]

Taking what both Llewellyn and Leiter state together, general themes emerge which form a basis of any claims about Realism. First, Realism is particularly interested in judicial decision-making. That is not its only focus but it does have an emphasis on that part of the law. Furthermore, Realists were specifically concerned with the ability to be able to predict judicial decisions. They wanted the ability to be able to know in advance the likely result of cases which came before the courts. Consequently, they have a specific focus on practical lawyering. In other words, the Realists' prime concern was not philosophical matters but legal ones. A result which was philosophically unsound would still have been a useful one, from their position, provided it was legally helpful.

In addition to this emphasis on practical lawyering, prediction and judicial decision-making, the Realists were also keen to explore the use of additional material (usually from social sciences) beyond the commonly accepted legal sources.[20] This had two implications.

14 Llewellyn (n. 11) 1238.

15 Ibid 1236–38.

16 Leiter (n 1) 21.

17 Ibid.

18 Ibid 26. The Determination Thesis is that 'choice of decision must, in fact, be sufficiently fettered so that prediction is possible'. The Generality Thesis is that 'these fetters upon choice must not be idiosyncratic facts about individual judges, but rather must be of sufficient generality or commonality to be both accessible and to admit the sorts of lawful generalizations that make prediction possible'.

19 Leiter (n 1) 30.

20 Not all of their efforts to increase the ability to accurately predict results relied, however, on the use of social science. For example, Llewellyn spent a considerable amount of time focusing on the difference between 'paper rules' and 'real rules': K Llewellyn, 'A Realistic Jurisprudence: The Next Step' (1930) 30 Columbia Law Review 447–53.

First, the Realists used these additional materials to aid and bolster the prediction enterprise. They thought that these additional materials could provide them with an ability to better predict the results of cases before the courts. They did not always agree on what materials were best placed to help in that regard but they did broaden the use of materials beyond those which had previously been used. Moreover, they found that additional materials provided further methods to evaluate and assess the law. This provided a benefit to the Realist movement because they wanted to explore not only the way the law worked but also how to make it work more effectively. The use of these broader materials allowed not only the prediction project to progress but also provided a needed constructive way to evaluate the law. It meant that the Realists had a much more inclusive idea as to what was relevant for law than previous jurisprudential theorists might have done.

One additional theme of Realism was the indeterminacy of legal rules. This theme developed from the broad themes discussed in the previous two paragraphs. While a focus on prediction or the use of social science research in exploring the law are methodological themes, the indeterminacy of legal rules is different. Instead of being about the methods used by the Realists, the idea that the standard written rules did not always determine the outcome in (at least some) concrete cases was a conclusion reached because of those approaches.[21] Even so, it is one of the themes of Realism which is most likely to get distorted in discussion. Some have argued that the claim attributed to the Realists is overstated. The claim made is that the Realists argued that the law was globally indeterminate. Leiter argues that the Realists' claims were only locally indeterminate and, in particular, limited to those cases which came before the courts.[22] Whether or not this is an accurate statement of the Realist position, the Realists did claim that cases were *pervasively* indeterminate even if not globally so.[23] A comparison helps to illustrate this point. H L A Hart argued that indeterminacy in law was only in cases of 'open texture'.[24] The law was subject to open texture because we are not clear about language. Words that we use have a 'core of certainty' and a 'penumbra of doubt'.[25] The 'core of certainty' consisted of those referents of a word which everyone agreed were contained within the definition. The 'penumbra of doubt' affected those cases where it was not clear if the word in question covered them or not. In his seminal example, Hart argued that a car would fit within the 'core of certainty' of the word 'vehicle'; a bicycle or skateboard, on the other hand, were less clear and covered by the 'penumbra of doubt'.[26] For Hart, these cases occurred because we were not always specific enough about our use of language.[27] In addition, we were bad at predicting the future and thus new information was rarely covered explicitly in a rule.[28]

The Realists did not believe that indeterminacy claims were so limited. They would have argued that Hart's 'open texture' was only part of the picture. Facts could often be indeterminate in the sense that they were filtered through witnesses' or judges' perceptions

21 Schauer (n 7). It should be noted the difference between particularism (the idea that a case is determined by factors unique to that case and therefore each decision is individualistic) and the view that factors outside the specified rules might determine the outcome in cases. The second view need not be particularistic if there is a pattern to those decisions even if that pattern involves factors outside of legal reasons as we generally conceive them. The Realists tended to take the latter view even if they are often believed to be proponents of particularism. See ibid.

22 Leiter (n 1).

23 Schauer (n 7).

24 H L A Hart, *The Concept of Law* (2nd edn, Clarendon Press 1994) 123, 128–36.

25 Ibid 123.

26 Ibid 128–29.

27 Ibid.

28 Ibid.

of importance and thus important information might be misremembered or valued differently by a judge.[29] There were also a number of rules which could potentially be in play in a particular case and the judges determined which rule applied and why it did so.[30] Additionally, a judge might read cases differently from her predecessors and apply rules in a different manner.[31] Finally, there might be cases where judges applied rules only after they had reached a decision on other grounds and the rules therefore only played a post hoc justificatory role.[32] To the Realists, judicial decision-making was a complex activity and the prediction of it included not only an understanding of the language used in the rule but of surrounding factors, all of which might affect the ability of the rule to determine the outcome in a concrete case.

Two examples highlight the distinction between the two positions. Those two cases are *Brown v Board of Education*,[33] the famous US Supreme Court case on the desegregation of schools, and *R v R*,[34] the UK criminal case which outlawed the marital rape exception.[35] Both cases are indeterminate as the rule set down in previous cases was not followed.[36] Indeed, both cases overruled the pre-existing legal rule – *Brown* overruled *Plessy v Ferguson*[37] and *R v R* overruled, among others, *R v Clarence*.[38] Neither decision involves what Hart referred to as open texture. *Brown* was covered by the 'separate but equal' doctrine set down in *Plessy*. *R v R* fits within the rule in *Clarence*. Nor is it a case where the facts in the subsequent cases (*Brown*, *R v R*) were not envisaged at the time of the previous cases (*Plessy*, *Clarence*). *R v R* is the same sort of case as *Clarence* on this issue.[39] While *Plessy* dealt with railroad cars and not education systems (as *Brown* did), educational systems were around during that time and the language of *Plessy* is broad enough to cover any societal arrangement.[40] They are not cases based upon our inability to see the future like

29 Frank (n 13).

30 K Llewellyn, 'Remarks on the Theory of Appellate Decision and the Rules or Canons about How Statutes Are to be Construed' (1950) 3 Vanderbilt Law Review 395.

31 K Llewellyn, *The Bramble Bush* (Oceana Publications 1951) 180.

32 Leiter (n 1) 30.

33 (1954) 347 US 483.

34 [1992] 1 AC 599.

35 The marital rape exception was a rule in UK criminal law that a wife, by virtue of getting married, had consented to all physical intercourse. It was therefore not legally possible for a husband to rape his wife: ibid 603–04. The genesis of the exception appears to be Sir Matthew Hale's *History of the Pleas of the Crown* (1736): ibid 604.

36 The same claim could be made about *Riggs v Palmer* 115 NY 506, 22 NE 188 (1889), the case famously used as an example by Ronald Dworkin (n 6) 15–20. The wills statute at issue in that case appears to be relatively clear on point. However, because the case is linked so readily to Dworkin's views on integrity and the law, other cases provide a better test for exploration for our purposes.

37 (1896) 163 US 537.

38 (1888) 22 QBD 23.

39 *R v R* (n 34) is certainly a clearer case of the marital rape exception. *Clarence* (n 38) dealt with a wife who had developed a sexually transmitted disease (STD) from intercourse with her husband who, while knowing he had the disease, did not inform his wife. The argument was that the fact that he had an STD but had not disclosed it vitiated the consent his wife had given to sexual relations. The court ruled that was not possible since, at least in part, she had consented to the marriage and that meant, necessarily, that she had consented to any and all sexual contact.

40 *Plessy v Ferguson* (n 37) 544. The Supreme Court specifically mentions schools noting: 'The most common instance of this is connected with the establishment of separate schools for white and colored children, which has been held to be a valid exercise of the legislative power even by courts of States where the political rights of the colored race have been longest and most earnestly enforced.': ibid. At several points, the Supreme Court uses the argument that, since it is acceptable in education to separate out the races (relying heavily on the fact that the DC School District was segregated), then it must be acceptable for train companies to do the same: ibid 544–45, 545, 551.

determining a case raising the question of whether a Segway fits into Hart's 1960s statute prohibiting vehicles in the park. Furthermore, in neither case does the court rely on language to change the applicable rule but overrule the previous decision.

What is particularly interesting about these two cases is that it would have been *possible* to make a case relying on Hartian open texture for the decision. The court in *Brown* could have argued that the Topeka School Board needed to bus black students to the 'white' school merely because they had failed to satisfy *Plessy*'s 'separate but equal' test.[41] In general, the facilities provided to black students did not fit under the standard definition of equal if equal means something beyond merely providing something within the same class to different groups.[42] They could have argued instead that equal meant that the facilities had to be of the same standard. Likewise, in *R v R*, the married couple had separated prior to the rape in question and the woman lived in a different house from the husband at the time of the rape.[43] The court could have argued that the marital rape exception did not apply under the circumstances as the separation had changed the applicable rule.[44] In neither case did the court in question take this way out. Instead, they argued that subsequent legal principles had made the previous rule unsound (if it had ever been sound) and the previous decisions were overruled.[45]

It is this sort of case that the Realists want to explore. Ambiguity of language is an important aspect of the law, but cases in which the court reaches a decision which does not follow clearly from previous decisions for other reasons are also worthy of consideration. To the Realists, this was all part of the same indeterminacy problem. In their thinking, there were a multitude of reasons why decisions might be unclear and we needed a comprehensive method for exploring the relevant issues. The Realists, unlike Hart, conceived of legal indeterminacy as being a broad practice involving a number of different factors all of which could potentially be applicable in a given case. They considered this all part of one process involved in judicial decision-making. So, Hart probably would have had a response to the issues created by *Brown* and *R v R*. They just would not have been covered by his conception of open texture. To the Realists, however, all of these sorts of issues were contained within the heading of legal indeterminacy.

In conclusion, we can state there are a number of important things we might find as indicative of the Realist position. They were particularly focused on practical lawyering as

41 Indeed, the lower court decisions in *Brown* involve four separate school districts (Topeka, Kansas; Clarendon County, South Carolina; Prince Edward County, Virginia; and New Castle County, Delaware) and there was an examination of whether those School Districts did provide 'equal' treatment: (n 33) 486. In the cases themselves, the Kansas schools were considered to be equal but those in South Carolina, Virginia and Delaware were not. According to Chief Justice Earl Warren (who wrote the opinion of the Court), equalisation measures had either been completed or were 'well on their way' by the time of the Court's decision: ibid 492.

42 As noted in n 41, the four school districts in question might have been exceptions to the rule. Nevertheless, anyone wanting a visual representation of the extent to which 'separate but equal' wasn't really equal need only look at pictures of, for example, water fountains in the segregated South during this period.

43 See (n 34) 614–15.

44 The Court of Appeal, it should be pointed out, does consider this possibility. It had been used in previous cases to outline exceptions based upon the issuance of a decree nisi, express arrangement by the couple, or by court order. The Court of Appeal, however, decides instead that 'the law should declare that a rapist remains a rapist subject to the criminal law, irrespective of his relationship with his victim': ibid 611. There is discussion by both the Court of Appeal and the House of Lords as to the word 'unlawfully' in the Sexual Offences (Amendment) Act 1976, s 1(1): ibid 610–11, 623. In particular, the question that arises is whether the Act of 1976 enshrines the marital rape exception in the law such that it cannot be removed by the Court. That issue, then, is somewhat subsidiary to the question as to whether or not the marital rape exception is a valid exception in common law.

45 *Brown* (n 33) 494–95; *R v R* (n 34) 623.

well as judicial decision-making. This included the ability to be able to predict the decisions that judges and courts might make in disputes which came before them. They also looked to expand the number of relevant sources of information which were considered both in relation to legal decision-making as well as to the process of evaluating legal decisions. This included a complex notion of what constituted legal indeterminacy which was grounded in the notion of a broad, pervasive conception of when the rules could be indeterminate and what the causes of that indeterminacy were.

3 Issues arising from Realism

There are two main issues which are brought up by the Realist examination of the law but are not resolved by it. The first question is what constitutes legal reasons for actions. In other words, what is encompassed by the word law? This is a standard question in legal philosophy but it takes on a special importance for the Realist project. Because of Realists' understanding of legal indeterminacy and their acknowledgment of the expanding class of available reasons for action, they cannot identify the distinction between legal and non-legal reasons by reference to the job performed by those reasons. Both types can provide reasons for legal decision-making. We could determine that something is a legal reason for action by looking to its source, but that also provides difficulties for the Realists. Since legal decisions are likely sources of law and the reasons for those decisions need not necessarily be law-related reasons, the Realists would need to explain how the action of a judge speaking or writing turned a non-legal reason into a legal one. Again, the easy answer (that the act of speaking or writing by a judge creates a legal reason for action) would not likely work. Not everything a judge speaks or writes would work as a legal reason (for example, a letter written to a friend). Even if it were limited to those times when a judge was specifically acting as a judge, this argument would not be particularly effective. If a judge stated that they had decided a particular decision by flipping a coin, the legal community would not see that as a legal reason for the decision even if the judge had, in fact, decided by that method. There must be some *further* criteria for distinguishing between legal reasons for actions and those non-legal reasons. Additionally, the Realists would be required to hold that there is a distinction between certain reasons called legal and those which are not because the use of post hoc justifications which they frequently commented on requires it. If all reasons are potentially legal reasons for action, then there is nothing wrong with using 'non-legal' reasons as justification. It is only if non-legal actions are seen as unjustifiable reasons for decisions that judges need to rely on so-called 'legal' ones.

This classification issue has bite because of the more overarching question of justification. Any legal system must provide reasons for individuals to abide by it if that legal system is going to continue to exist. We must accept those decisions made by those acting under legal authority as binding. This causes particular problems for Realists because of the varied influences on legal decision-making. There must be an explanation why a decision taken by a judge based on extra-legal factors provides a sufficient reason for holding that (1) we are required to follow that decision, and (2) society is entitled to believe that we will follow the decision. The argument in support of the legitimacy of law is easy to make in situations where the grounds for our legal decisions are limited to legal reasons for action. If judges could only make decisions based upon previous case-law and statutes, authority for such decisions is on the basis of the existing rules – for example, that they offer fair warning of likely legal consequences, that they have been subject to democratic review or that they are authorised by a social contract. However, if the Realist critique is accurate, then decisions are being made for reasons which have not been previously agreed, are not vetted in the public sphere and are, at least sometimes, hidden from view. Such decisions are ones

we would consider suspect in terms of legitimacy and justification. Since these decisions are extra-legal in that they do not rely on legal reasons for action, then even if the overall legal system is justified, these particular decisions may not be.

We can explore this by way of example. Suppose we had Holmes' famous 'bad man'.[46] Let us further suppose that our bad man lives in the community covered by Hart's 'no vehicles in the park'.[47] Our bad man wishes to bring a bicycle into the park but does not want to suffer any adverse legal consequences for doing so. Bicycles have been allowed in the park before. However, they were children's bicycles and our bad man wants to bring in a heavy-duty mountain bike so he can train for an upcoming race. The bad man examines all of the information related to the regulation and decides that, since bicycles have been allowed in the past, he is likely to be allowed to use his bicycle as well. To his surprise, when he brings his bicycle into the park, he is cited for violating the regulation. When he contests this decision in court, he is told by the judge that he cannot bring a bicycle into the park as it is a vehicle. In particular, the judge relies on safety arguments to argue that the legislative body in question meant to include bicycles within the definition of vehicle. This is over the protestations of the bad man who, having looked through the law in question and the cases interpreting it, asserts that no previous decision has made any mention of 'safety' when determining the definition of vehicle under the rule.

Under this scenario, why should our bad man feel that the judge's decision is a legitimate legal decision? We could fall back on the argument that if he fails to abide by the decision, he is subject to sanctions (e.g. a fine) and if he fails to comply with the sanctions, further more powerful sanctions might be applied (e.g. incarceration). That does not necessarily provide a justification for the decision; it merely coerces our bad man to accept it. If we do not rely on coercion (for example, if we want to argue that the decision should be considered right or proper), we need to look somewhere else. Nor would it be useful to argue that the Realists did not concern themselves overmuch with this justification problem.[48] Even if they did not care about legitimacy and justification, nothing prevents us from asking the question in relation to their theory.

Additionally, if the Realist critique is accurate, we need some justification for the system as a whole or at least the way the system is run. Why would we design a system which allows judges to make decisions based upon non-legal factors? Could we not develop some sort of mechanistic system for determining answers in cases other than Hart's open texture ones? Could we not require that our legislatures be more specific when developing statutes? Alternatively, could we not construct a rule when anything which is not specifically prohibited is excluded from regulation? To put the questions another way, why do we persist with the legal system we have? Why do we put such power into the hands of a few individuals, often ones who are neither representative of the population at large nor democratically elected? In order to justify the legal system as a whole we need some semblance of an answer to these sorts of questions. Moreover, these questions persist even if we are only interested in a purely descriptive account of the law. We must still have some understanding as to why individuals within a system feel bound by it. We must understand why it is that they feel an obligation (using Hart's terminology)[49] to obey the law despite

46 O W Holmes, 'The Path of the Law' (1897) 10 Harvard Law Review 457, 460–61.

47 Hart (n 24) 128–29.

48 Leiter (n 1) 114–15.

49 Hart (n 24) 6–7.

misgivings we might have about the way the system operates.[50] We must also have an explanation for why individuals continue to abide by the law even when it is against their interests to do so. We must further explain why judges, when acting in their official capacity, act as if they can provide justifications for their actions and expect that these will be accepted as right and proper ones. We must understand why it is that judges continue to couch their decisions in legal language and legal reasons even if, at least some of the time, the answer they give is not truly why they decided as they did.

Realism does not really give us answers to these questions. The Realists provide us with a reliable description of what is occurring in legal decision-making but cannot provide us with an answer as to why that ought to be the case. Even when they provide optimal ways of working within the system they do not tell us why the legal system needs to be in such a state as to require those optimal ways in the first place. We need to look elsewhere.

4 Dworkin as an answer to Realist issues

If Realism cannot provide us with an answer to these important questions, another theory can provide it without requiring that we abandon the conclusions from Realism that we wish to keep. Those looking for an answer to justificatory questions while still maintaining the basic tenets of Realism can look to Dworkin's law as integrity to provide them. In other words, Dworkin's theory is not incompatible with Realism.

It is worth beginning that discussion by examining the existing parallels between Realism and Dworkin's work. One fundamental parallel between the two is that both focus on judicial decision-making. The Realists, as mentioned above, paid particular attention to the role of judges, both at the trial court and at the appellate level. Dworkin, with the focus on law as integrity, Hercules and the right answer thesis,[51] also focuses on the judicial aspect of law.[52] This does not mean they exclusively focus on judicial decisions because neither does, but the judicial aspect of law is more important in their theories of law than, for example, it is in Austin, Kelsen or even Hart. Moreover, both focus on practical lawyering. Dworkin gives a special prominence to exploring the question of legal disputes and disagreement.[53] This emphasises the question of how it is that legal professionals can disagree about the fundamental claims of a particular legal system.[54] The question in its simplest form is a pragmatic one – how do we explain legal disagreement? Dworkin and the Realists pay attention to such questions even if other legal theorists like Austin, Kelsen and Hart do not. Additionally, Dworkin and the Realists often use case-law to provide examples of particular claims.[55] Their purpose in doing so is to evaluate not only the theoretical basis of law, but to actually test whether that basis is correct. Both think they are providing a method by

50 We might, indeed, even question whether or not people actually feel an obligation to obey the law. This is presumably an empirical question (or could be an empirical question) but I am not aware of anyone who has ever explored it empirically. I am grateful to Frederick Schauer for bringing this point to my attention.

51 The extent to which Dworkin still maintains the relevant thesis as a 'right answer' thesis as opposed to a 'best answer' thesis is a bit murky. Many contend that Dworkin now talks about 'best answers' as opposed to 'right answers'. However, the most recent complete statement of his views on law (*Justice in Robes* (Belknap Press 2006)) still maintains at least some focus on 'right' answers and a 2006 book which explores his views makes repeated references to right answers, a characterisation he does not refute in his response: S Hershovitz (ed), *Exploring Law's Empire: The Jurisprudence of Ronald Dworkin* (Oxford University Press 2006). As best I can tell, then, Dworkin's view is still largely similar on point to the one he has always professed and any distinction between a 'right' answer and a 'best' one is semantic. Even if this is not true, I am not sure it changes any of the analysis in this article.

52 Dworkin (n 6) 14–15.

53 Ibid 6–7.

54 Ibid 3–6.

55 Ibid 15–30.

which judges can explore their own legal decision-making, or to evaluate the real life decisions that judges make. Furthermore, both seek to expand the number of reasons available for evaluating the decisions of legal decision-makers, although the types of reasons are not the same. While the Realists looked to social science data, personal idiosyncrasies of the judges or other similar factors, Dworkin has focused his attention on principles.[56] Thus, both look to increase the number of reasons found to explain legal decisions beyond legal rules, whether derived from statute or previous court decision.

These parallels between Dworkin's theory of integrity and Realism show the possibility of cross-germination between the two theories. However, it might be argued that these similarities are insufficient on their own and there is the very real question of legal indeterminacy to consider. While the Realists were keen to argue that the law was at least partially indeterminate, Dworkin has maintained that legal principles reduce the indeterminacy of law.[57] Some might argue that, despite any initial facial likenesses between the two theories, the essential claim about indeterminacy creates an insurmountable barrier between the two. A deeper analysis reveals this is not the case. While Dworkin does argue that integrity minimises (if not completely eliminates) a multitude of answers to a particular legal question, this is not because of a fundamental disagreement with the Realist project as such. Dworkin must initially hold something close to the Realist conclusion on indeterminacy for the argument for principles to have any meaning. At the very least, he must start from a position much closer to Realism than it is to Hart's notion of open texture.

In order to understand this, we must explore why Dworkin thinks principles are so important. For Dworkin, legal rules cannot account for a number of cases. They are too inflexible and too all or nothing to be able to adequately explain the complexity of legal decision-making. They cannot, for example, explain the decision of the US Supreme Court in *Brown*, one of the decisions he notes at the beginning of *Law's Empire*.[58] Dworkin likewise sees 'hard cases' as being a more comprehensive phenomenon than Hart ever admits.[59] For Hart, they are a small number of decisions;[60] Dworkin, on the other hand, argues that hard cases are frequent in the law.[61] Dworkin argues that there are similarities between lines of cases which are used to bring disparate legal issues together in order to provide a means for resolving these cases. Those connections are what he calls principles. *Brown*, consequently, can be explained as a decision about something larger than the inflexible legal rules associated with previous cases like *Plessy*. It is a discussion of equality, a connection between a number of different sorts of cases so that the decision in *Brown* fits into the larger legal system.[62] All of these disparate cases form a coherent system by being mutually supportive of each other.

Principles, then, are broad statements which express moral beliefs within a particular community.[63] They apply across a range of cases and interact with other principles as well as legal rules in order to provide the most convincing legal arrangement possible under the circumstances.[64] This idea of principles is much more expansive than is necessary to deal

56 R Dworkin, *Taking Rights Seriously* (Gerald Duckworth & Co 1977).

57 Dworkin (n 6).

58 Ibid 29–30.

59 For example, he states that any case from any case textbook would provide examples of the principles he is talking about. See e.g. Dworkin (n 56) 23.

60 Hart (n 24) 154.

61 Dworkin (n 56) 81–130.

62 Dworkin (n 6) 379–92.

63 Dworkin (n 56) 82.

64 Ibid 81–130; Dworkin (n 6) 95–96.

with uncertainty in a Hartian 'open texture' sense. If the only type of indeterminacy was based upon the penumbra of doubt, it is unnecessary to consider broad principles which apply to a large range of cases and bind the system together. There would be no need to go through that process in a majority of cases so the principles themselves would be superfluous. However, if the indeterminacy that Dworkin is considering is the pervasive one used by the Realists, then principles become a more reasonable idea. If almost any case might be subject to indeterminacy based upon the facts, a disparate collection of rules, or other factors, then moral principles which bind the legal system into a coherent whole become necessary. In other words, there is no reason to have a broad principle to explain whether a bicycle is a vehicle or not under a simple regulation prohibiting vehicles in the park. We do need principles to explain decisions such as *Brown* and *R v R*. Dworkin must start from a position that the pervasive indeterminacy of the Realists is a more accurate assessment of legal decision-making than the more limited notion of 'open texture' presented by Hart since it is only a more pervasive and comprehensive indeterminacy which requires the use of principles in the way that Dworkin uses them. The pervasive legal indeterminacy of the Realists is not therefore an anathema to the Dworkinian use of principles, it is a precursor to it. The Realists' ideas about indeterminacy provide support for an important presumption Dworkin makes about the law. As a result, it does not adversely affect the ability to see connections between the two sets of theories but supplements the similarities which have already been discussed.

With this understanding, it becomes possible to explain how Dworkin's law as integrity can plug the justification gap left by the Realist critique of law. That way is the right answer thesis that Dworkin provides.[65] To explain this, let us return to our bad man wanting to bring a bicycle into the park. Our bad man brought a bicycle into a park covered by a rule which stated that 'no vehicles were allowed in the park'. Having explored the statute and previous case-law, our bad man had determined that there was no specific part of the legislation which prohibited bicycles per se nor was there any case-law which prohibited bicycles. Despite this, when the case came before a judge, it was determined by that judge that bicycles were included within the definition of 'vehicle' such that bringing one into the park constituted a violation of the law. Let us suppose, however, that the jurisdiction in question is slightly different from ours in terms of procedure. The law is exactly the same but litigants have the possibility after the court has made a decision to question it. Once a decision has been handed down, the judge must discuss that decision with the parties in question should they wish to do so. Our bad man requests such a discussion with the judge and asks why the court ruled the way it did despite the fact that there was no specific mention of prohibiting bicycles either in the legislation itself or in the previous case-law. Our judge has several responses available to her. She could reiterate the safety concerns which formed the essential part of the judgment itself. This will not provide much help. Our bad man will ask where these safety concerns came from since they did not appear in either the statute or previous case-law. What our judge will need to do is to provide a reason as to why those safety concerns could be used to make a decision, even if they were not specifically noted in the law before this case.

Note that the use of the Hartian response to our bad man will not be an effective mechanism. If our judge argues that there is a penumbra of doubt about the word vehicle and that she exercised her discretion in order to determine that 'bicycle' was covered, this will lead to serious problems with our bad man. He will argue that this exceeded her authority (since she should only act on the basis of the law, not on her 'discretion'). Even if

65 Dworkin (n 51) 41-43. Again, I do not think this analysis changes if we consider Dworkin to now hold a 'best' answer thesis as opposed to a 'right' answer one. See n 51.

he does not do that, he will ask why she exercised her discretion in this particular way. Why, he might ask, did she exercise her discretion in order to find him guilty of violating the law instead of using her discretion in order to find him not-guilty? Our judge needs an answer to this question in order to satisfy the bad man and one that does not circle back to the safety concerns. If our judge answers 'Because I thought the safety concerns were important,' this only leads back to the same questions as to why the judge was entitled to consider them in the first place. Alternatively, the judge could fall back on claims of authority. She could claim that she was entitled to decide the way she did because that is the job she has been given by our legal system. The decision is authorised by the position she holds. That also would not find the support of our bad man. He will argue in response that the mere fact that her position gives her the authority to make decisions in the legal system does not mean that authority is unfettered. She would not, for example, be able to sentence him to death or public flogging for breaking the statute in question. Such a decision on her part would be an abuse of power. Why then, is her decision not a similar abuse of power, again because there is no specific indication that bicycles are covered? In other words, even if our bad man accepts that her role as judge provides her with the authority to make decisions generally, it does not follow that it justifies *this* decision in *this* case. Moreover, neither argument will necessarily prevent our bad man from bringing his bicycle into the park on subsequent occasions. Both justifications – the Hartian open texture one or the one based upon her authority as judge – can be seen as being personal to the specific judge. Our bad man, then, may decide in subsequent cases to try his luck with different judges.[66] While our first judge might decide that bicycles are covered, the second (and third, etc.) may not.

If our judge wants to assuage the bad man, she needs to convince him that the decision she made was the correct one, despite what he perceives to be a lack of grounding for it. In order to do that, our judge is going to have to provide evidence that safety concerns are valid even if they are not specifically stated in the statute or covered by previous case-law. One way to do that will be to show our bad man that other courts have used safety concerns before in a range of cases, even where there was not a previous statement about safety in case-law or statute. While that is an important part, it is not the only part. She will further need to explain why this case is similar enough to other cases where safety has been an issue for it to be an acceptable reason for the decision here. She would further have to argue why, even if there are alternative possible justifications for decisions, the one she has chosen is the best one to have used under the circumstances. Notice that she cannot simply claim that her decision is one of a range of possible alternatives. She cannot simply claim that her decision is based upon a reasoned argument but that there were alternatives which were just as likely. Our bad man will simply insist she justify why she did not choose one of those other alternatives which would allow him to bring his bicycle into the park. She must claim that her decision is right or correct or best under the circumstances. That does not mean that she cannot accept that our bad man has made a reasoned and reasonable argument. She must have to claim, however, that hers is better than all alternatives. Moreover, she needs to claim that her decision is not just right or best in her opinion but right objectively. Her claim also has to be stronger than a claim of preference on her part for the argument she has given. If she merely claims a preference, then our bad man will simply query why her

66 This is the concern with particularism in relation to judicial decisions. If we cannot presume that the decisions have some sort of general application, then there is nothing to prevent an individual from attempting an action which brought a sanction in an initial case to try again. This point will be considered more fully in Section 5.

preference is better than his – why it is entitled to greater weight.[67] If she cannot, then it simply becomes a claim about taste much like someone's preference for baseball over football. If she wishes to justify her decision, she needs to be able to state to the bad man that her decision is best not only for her, but for everyone. It must be an objectively true statement of the law.

What has been described, though, is no different than Dworkin's right answer thesis. He makes the following statement of the right answer thesis:

> It is a claim made within legal practice rather than at some supposedly removed, external philosophical level. I ask whether, in the ordinary sense in which lawyers might say this, it is ever sound or correct or accurate to say, about some hard case, that the law, properly interpreted, is for the plaintiff (or for the defendant.).[68]

He further states:

> Have you yourself found any ordinary legal argument on balance the soundest, in any kind of hard case? Then you, too, have rejected the no-right-answer thesis I take to be the target of my own claim.[69]

Dworkin's thesis requires that judges be able to make decisions which are objectively right, which is what our judge has attempted to do in this case. Her decision, if she wants to be able to justify it to our bad man, must be a decision which she can claim to be right and to be right objectively. She must be able to claim this despite the fact that others might disagree. In other words, she can accept that others might come to a different conclusion but that those conclusions are not as good as hers despite the fact that they may be reasoned and reasonable arguments. None of this, however, requires a refutation of the indeterminacy thesis as the Realists conceived of it. In fact, it requires the indeterminacy thesis as a starting point. It must start from the idea that cases are indeterminate in the pervasive way the Realists discussed in order for the right answer thesis to have sufficient bite.

Additionally, this reasoning works in the more complex cases of judicial decision-making used in *Brown* and *R v R*. In both cases, the side which ultimately lost (the School District of Topeka, Kansas, in *Brown* and the husband in *R v R*) would have, given the opportunity, questioned the decisions given in their respective cases. Indeed, in their cases, they probably have a stronger argument than our bad man since there was actual case-law in their favour prior to their actions. In order to justify those decisions, the judges would have to use the right answer thesis to convince them. They would need to argue that the previous cases, despite being legal decisions which argued that their actions were acceptable, were not the best statement of the law. In order to do this, they would need to show that higher values (e.g. equality under the law) were the most crucial values at stake and those values required the decision that the judges made. They would have to argue that these were the correct decisions no matter which judge made them even though previous judges had, in effect, gotten it wrong.[70]

---

67 It is possible that our judge will return to the argument about authority. However, the same problems would exist here as it does with the use of it earlier. Our bad man would argue that, even if she has general authority, it only exists in situations where she is acting appropriately. If she is exceeding her authority (as he would claim she is doing here), then it provides no greater justification for her decision now than it did before.

68 Dworkin (n 51) 41.

69 Ibid 42.

70 In making this claim, I do not want to forestall the idea that principles can change over time. It may be that a principle which was at one time valid has ceased to have validity and therefore decisions based upon that principle in the past are likewise invalid. That change, though, may be attributed to the creation of new, more important principles (or at least more important at the present time) and we would need a way to explain this. Moreover, even if principles lose their validity, we need some method for being able to explain to members

Dworkin's theory of integrity, then, fills the justification gap which exists in the Realist explanation of the law without forcing one to refute many of the central claims of Realism. It also provides further benefits. First, it can help explain cases which the Realists have often had trouble with – those cases where the judge appears compelled to reach a decision despite being against it. Realist approaches to judicial decision-making suggest that judges often have enough flexibility to avoid reaching decisions which they find to be wrong. They ought to be able to manipulate precedents, the language in statutes, etc., so that they can reach a conclusion which they find satisfactory. That is not always the case. Sometimes judges reach decisions which they feel are compelled by previous case-law or statute even though they do not believe they are the right decisions on some other metric (e.g. morally, politically, or on policy grounds). Dworkin's theory can provide us with an answer to that question. Judges in these cases feel bound to reach the best decision in light of the previous law. Even if they personally disagree with the decision, a judge can come to a conclusion that a particular interpretation is the best legal answer in a specific case. One can decide that a particular answer is the best one, even if it is not personally a satisfactory one.[71]

5 Conclusions

In conclusion, there are still significant differences between law as integrity and Legal Realism. There are also a number of similarities and one need not give up all of the central tenets of either theory to accept that the other theory has benefits. For at least a couple of reasons, this answer might not be surprising. Dworkin does treat pragmatism less harshly than conventionalism. He claims that it 'so far from fitting our legal practices worse than conventionalism does, fits them better'.[72] More relevant is Karl Llewellyn's approach to judicial decision-making which he terms 'Grand Style'.[73] For Llewellyn, there were two types of judicial decision-making. The first, formal style, was a mechanistic way of making decisions involving only the logical extension of previous cases and nothing further – no further understanding of policy, principle or broader values.[74] Grand Style, however, was expansive. It considered a broader range of materials including important issues of

[*note 70 continued*] of society why a principle, which was at one time valid, no longer is. So, even if the judge does not make the claim that previous judges made a wrong decision in a case, they must be able to explain why that decision would not be right now.

71 To provide a personal (admittedly anecdotal) example, I once worked for a trial court judge in Pennsylvania as a law clerk. One of the cases brought before the judge was a contract dispute involving the construction of a garage. The homeowners had contracted with a construction company for the construction of the garage to their specifications but were unhappy with the result and wished not to pay the amount owed. While their claims in contract had failed, their lawyer also relied on the Consumer Protection Law in Pennsylvania at the time. That law required that any sale involving a 'contact at the home' required that the purchasing party receive notice they had the right to rescind the contract within three days. If the notification was not provided at the time the initial contract was signed (which it is was not in this case), then the notification period began to run once the purchasing party became aware of the right to rescind (i.e. when their lawyer told them about it). 'Contact at the home' had been determined earlier by the appellate courts to include cases such as construction ones despite the purpose of the law being to deal with high-pressure door-to-door sales. Utilising the Consumer Protection Law, the homeowners then rescinded the contract. There was a provision within the statute that required the return of the items purchased but only if the selling party requested them within a set time period. Because the items in question were incapable of being returned, the attorney for the contractor did not request them back. At oral argument, the lawyer for the homeowners admitted that, if the materials had been requested, they would have been unable to provide them and the consumer protection ground would have failed. However, since they were not, they were simply not required to return them. While the judge I worked for thought the decision reached a dreadful result, the law did appear to require that he rule in favour of the homeowners and allow them to rescind the contract without having to 'return' the garage.

72 Dworkin (n 6) 157.

73 K Llewellyn, *The Common Law Tradition: Deciding Appeals* (Little, Brown & Co 1960) 36–37.

74 Ibid 38.

principle.[75] It further focused on the creation of good decisions on a large scale, even at the expense of static legal doctrine.[76] The broad interpretive nature of Llewellyn's Grand Style is thus a forerunner to the interpretive method favoured by Dworkin.[77] Llewellyn also often makes statements which appear similar to Dworkin's view on rules and principles. For example, in *The Theory of Rules*, Llewellyn argues for the use of the purposes of law in order to provide a more coherent basis for the law.[78]

He also prioritises the use of conceptions of justice and fairness, two of the foundational principles Dworkin deems important in the law.[79] Llewellyn's view, thus, looks remarkably similar to Dworkin's views on integrity. He highlights the purposes of reasons as opposed to slavish dependence on written rules; the notion of change according to those purposes and reasons for rules and that it is these elements in the law, not written rules, which provide for the stability of the law.[80] Llewellyn, consequently, appears to have a number of similar views to those expressed by Dworkin.

The real issue is what this might mean for our understanding of both Realism and law as integrity. One initial conclusion we can reach is that neither theory is quite as extreme as its detractors suggest. Realism has long been a frequent whipping boy for legal theorists because it is believed to be an over-exaggerated farcical theory about judges which lacks sufficient intellectual rigour.[81] Dworkin has also been subject to criticism, particularly in relation to the right answer thesis as it too is considered an over-exaggerated position.[82] However, neither theory need be understood in the extreme view. The acceptance of the Dworkinian view does not mean that one need accept that legal problems have solutions which are discernible like mathematical problems any more than one need accept that judicial decisions are based upon what the judge had for breakfast in order to accept Realism. Both positions are much more moderate than they may initially seem. Moreover, both focus on the important aspect of problem cases. While other theories have attempted to start with the areas in which there is convergence, both Realism and Dworkin appear to begin their analysis with cases of disagreement. Dworkin and Realism both seek to explain how it is possible for disagreement to exist in law instead of trying to explain our general agreement about the law. Realism and Dworkin provide a much more satisfactory view of those types of decisions as a consequence. We can further take from this the view that both theories see law as being part of a community. The law is not a separate institution which acts on its own without any input from the other important aspects of our lives. Law is

75 As Schauer notes, Llewellyn 'often stressed the role of the judge in seeking to reach, albeit in small steps, the best solution to a general social problem': Schauer (n 7). See also W Twining, *Karl Llewellyn and the Realist Movement* (Weidenfeld & Nicolson 1973) 203–69.

76 Llewellyn (n 73) 36–37.

77 For a similar view in relation to Llewellyn's view of principles, see M D A Freeman, *Lloyd's Introduction to Jurisprudence* (8th edn, Sweet &Maxwell 2008) 1016.

78 K Llewellyn, *The Theory of Rules*, F Schauer (ed) (University of Chicago Press 2011) 136–37. It is worth noting that, while Llewellyn argues against the use of 'principles' in this passage, his understanding of 'principle' is different to Dworkin's. He objects to principles used when 'it is to serve, and can claim to serve, as a basis for judgment not only independently of any countervailing authoritative and more explicit words, *but also independently of any examination of its own underlying reason*': ibid 136 (emphasis in original). For Dworkin, of course, any principle requires understanding an application of the underlying reason for the principle. It thus corresponds to what Llewellyn wants to be done with reasons – 'it requires by its formulation and in its formulation to give reason for it to be taken as *the* reason, and to persuade that there *is* reason for it to be taken as *the* reason': ibid 137 (emphasis in original).

79 Twining (n 75) 215.

80 Llewellyn (n 78) 142.

81 Leiter (n 1) 1.

82 Dworkin (n 51) 41-43.

important because it incorporates, uses, evaluates and modifies the society and communities in which we live. We cannot explore and explain the law without coming to some understanding about what it is supposed to effect, both in terms of real life implications empirically as well as the underlying principles of the society in question. Law, then, is not a closed set, but one which necessarily involves a range of different factors.

Additionally, exploring how Realism and Dworkin's integrity can work together sharpens focus on the important elements involved in judicial decision-making. As has been noted above, Realism often suffers from criticisms about particularism. It is alleged that because the Realists viewed legal decisions as indeterminate, that meant they must believe that each individual decision was particular to the factual scenario. This, though, is a mischaracterisation and one that Dworkin helps to combat. If one sees indeterminism as a necessary ingredient of, first, Llewellyn's views on judging and, then, Dworkin's idea of law as integrity, it becomes clearer why indeterminism does not lead, necessarily, to claims about particularism. Understanding that the set of factual, legal and moral interactions which go into a judicial decisions are complex does not mean that prediction is not possible nor does it mean that general rules cannot be created from those decisions. Instead, it means that, in order to properly understand the reasoning behind those decisions, it is crucial to be able to grasp the entire picture. Some of those might be particularities about a particular judge – Lord Denning's noted bias towards the plight of elderly ladies is one example – but most involve aspects of judicial decisions which are normalised within the practice itself. These might be notions of fairness or justice which are outside of the strict legal rules; these might be biases or preconceptions which are implicit within the legal or more general community in which a specific court is situated.

Both Dworkin and the Realists, then, help us focus on what is really at the heart of problems about judicial decision-making. Decisions made by judges are difficult to understand jurisprudentially because of the wide power and scope implicit in the process. Judges are not (and indeed probably cannot) be bound by a mechanistic application of pre-existing rules. If they could, we would have little reason to have people engaged in the practice instead of sophisticated computers which could sift through the legal statutes in an attempt to determine the most efficient legal outcome. That does not mean that we expect judges to be totally unfettered. What we expect is a balance between the strict legal rules and the flexibility needed in specific cases. Using Dworkin's theory of integrity and the Realist indeterminacy thesis allows us to explore in greater depth that balance.

Dworkin's theory of integrity can also help shore up the prediction programme that the Realists were so interested in. If Dworkin's view of the judicial decision-making is accurate, then the use of principles should make predicting what a judge will decide easier. If we know, for example, that judges will use safety measures because that is a significant concern within a society, then we would be able to predict that heavy-duty mountain bikes are unlikely to be brought into a park unless adequate precautions are taken. This will allow us to predict the use of these measures in cases, even if past case-law or statute does not specifically indicate their use. Prediction, as the Realists wanted, becomes a more systematic and easier process as a result.

Furthermore, if we understand where there are similarities we can then begin to focus on those differences which are truly at issue. There is still the considerable issue of determining the extent to which reasons for decisions which are not previous cases or statutes are considered legal.[83] For Dworkin, reasons for actions which are used by judges (i.e. moral principles) are legal reasons. Indeed, Dworkin states that a judge, should she wish

83 Schauer (n 7).

to be able to provide real justification for a decision, must use reasons which are legal. It is not possible for a judge to use a non-legal reason, or at least not use one and still provide a justifiable legal decision. Conversely, the Realists did not consider the additional reasons for action that judges used to be legal. Exploring this difference between the two theories might help us better understand the extent to which we must rely on 'legal' reasons for decision-making as well as the extent to which we can fruitfully have a single definition of 'legal'. It might be, instead, that we would be better off looking at a multi-faceted description of legal (for example, something closer to what Llewellyn did with his idea of 'law jobs')[84] than we are focusing on a single overriding definition.

Another issue which is worthy of further exploration is the distinction between justification and coercion. I have argued that Dworkin follows Realism and helps to close the justification gap that exists in that older theory. This is important only if one believes that the justification gap problem is important and there is nothing that requires a legal theorist to believe that it is. One can simply rely on the notions of power and coercion to explain why individuals follow the law. Such a view, however, can be impoverished. Austin attempted to explain the law as a function of coercion and power and most modern jurisprudential scholars believe his theory provides a distorted view of the law. The common view now is that law is much too nuanced and complex to be explained only in terms of coercion. Law provides not only a coercive element but a facilitative one as well and legal theories need to be able to analyse these important elements. Exploring the differences between Realism and Dworkin on this important issue can provide a mechanism for the larger issue of justification within the law.

Not all problems, then, are eliminated even if we look to explore similarities between theories as opposed to differences. What we can take from this overall is the view that legal theory need not be an all or nothing thing. We need not accept wholeheartedly only one theory and argue that all others are incorrect. This is rarely accurate and usually can only be explained by mistakenly describing the theories of others. In reality, there is often far more in common between legal theories than there are differences. If we examine theories in order to properly understand those things which a theory might do well, then we are likely to be better able to explain the complex phenomenon that is law and legal systems generally. This does not mean, of course, that we can eliminate all differences between legal theories. There will still be disagreements, even on important fundamental issues. Nor does it mean we ought to give up critiquing jurisprudential theories. Of course we should not. Exploring similarities to the extent we explore differences, however, can be a more useful endeavour in at least some circumstances. Our critiques, then, ought to look at ways in which theories correspond to each other just as much as they explore ways they differ.

84 K Llewellyn, 'My Philosophy of Law' in Julius Rosenthal Foundation for General Law, Northwestern University, *My Philosophy of Law: Credos of Sixteen American Scholars 1941* (Boston Law Book Co 1941) 183–97, 185–86.