

NORTHERN IRELAND LEGAL QUARTERLY

Making a Difference? Why We Need a More Diverse
Judiciary (Lady Hale)

The Confidentiality of Medical Records:
Informational Autonomy, Patient Privacy, and the
Law (Joan Loughrey)

Beyond Rationality? Judicial Review and Public
Concern in the EU and the WTO (Chris Hilson)

Belfast Leases, Lord Donegall, and the Incumbered
Estates Act, 1849 (C.E.B. Brett)

Board Diversity: Can Sex Discrimination Law Help?
(Alice Belcher)

A Critique on Legal Analysis of Local Government
and the Central-Local Relationship
(Barbara Mauthe)

The Church of England: Some Historical Reflections
on a Constitutional Conundrum (Charlotte Smith)

Resolving the Duress Dilemma: Guidance from the House
Of Lords (Ray Ryan and Des Ryan)

Full table of contents inside

Vol. 56 No.3

Autumn 2005

CONTENTS

Making a Difference? Why We Need a More Diverse Judiciary
(Lady Hale) 281

The Confidentiality of Medical Records: Informational Autonomy,
Patient Privacy, and the Law *(Joan Loughrey)* 293

Beyond Rationality? Judicial Review and Public Concern in the EU and
the WTO *(Chris Hilson)* 320

Belfast Leases, Lord Donegall, and the Incumbered Estates Act, 1849
(C.E.B. Brett) 342

Board Diversity: Can Sex Discrimination Law Help?
(Alice Belcher) 356

A Critique on Legal Analysis of Local Government and the Central-Local
Relationship *(Barbara Mauthe)* 373

The Church of England: Some Historical Reflections on a Constitutional
Conundrum *(Charlotte Smith)* 394

Resolving the Duress Dilemma: Guidance from the House Of Lords
(Ray Ryan and Des Ryan) 421

BOOK REVIEW

Devolution, Law Making and The Constitution: Edited by Robert Hazell
and Richard Rawlings 431

Published four times yearly by SLS Legal Publications (NI), School of Law,
Queen's University Belfast, Belfast BT7 1NN, Northern Ireland.

ISSN 0029-3105



MAKING A DIFFERENCE? WHY WE NEED A MORE DIVERSE JUDICIARY

Lady Hale

*Law Centre (Northern Ireland) and SLS Lecture 2005**

“I have nothing against innovation but this has never been done before”

On this basis, or something very like it, women were excluded from the legal profession until the Sex Disqualification (Removal) Act 1919. But it is now 86 years since women were allowed to join the legal profession and aspire in due course to become judges. It is more than half a century since the first women sat as professional judges in our courts, The very first was Sybil Campbell, who was appointed a metropolitan stipendiary magistrate in 1945. She is nowhere near as famous as Dame Elizabeth Lane, the first woman county court and later High Court judge, and Dame Rose Heilbron, the first woman Recorder and second High Court judge. It is nearly twenty years since women began joining our law schools in equal numbers with men. Yet our judiciary is still predominantly male and pale.

In England and Wales, less than 11% of our salaried (permanent) judges in the ordinary courts are women. Fewer still belong to an ethnic minority (the published figures do not give the full picture, perhaps because not everyone is happy to have their ethnicity recorded). Your figures in Northern Ireland are rather better, with over 17% of your salaried judges in the ordinary courts being women. But percentages can be misleading when numbers are small. In both our countries, the situation is worse the higher one climbs in the judicial ranks. You have no women in the High Court or above until you get to me.¹ Only 10 of our 106 High Court judges are women; one of them is the very first from a black or Asian ethnic minority. Only 2 of our 37 Court of Appeal judges are women, none from an ethnic minority. Since the retirement of Dame Elizabeth Butler-Sloss at the end of last term, none of the four Heads of Division – the Lord Chief Justice, Master of the Rolls, President of the Family Division and Vice-Chancellor – is a woman. And it took until last year for a woman to be appointed to the top court in the United Kingdom – the appellate committee of the House of Lords.

The proportion of women amongst the fee paid (part time) judiciary in both our countries is higher, over 20%, and some find this an encouraging sign of better things to come. But even if all these women go on to become permanent judges, it will take a very long time indeed before the proportion of women on the bench mirrors the proportion of women entering the legal profession. It would be unfair to describe our pale male judiciary also as stale – but it is drawn from a comparatively narrow range of social and

* Given at Queen’s University on 17 June 2005.

¹ Lady Hale is the first woman to be appointed to the appellate committee of the House of Lords (editor’s note).

educational backgrounds and from a very narrow range of legal professional backgrounds. I am doubly unusual, not only in being a woman, but also in having made my professional name as an academic and public servant rather than as an advocate.

I believe that in all three parts of the United Kingdom, it was largely concern about the lack of diversity – on a variety of measures – in the judiciary which led to calls for reform in the judicial appointments system. This is not to decry the efforts made by Lord Chancellors and their departments to make improvements. But things have really started to move since we have had a Commissioner for Judicial Appointments, both in Northern Ireland and in England and Wales, and Scotland has led the way in having a Judicial Appointments Board. Now you have your own Judicial Appointments Commission, launched this week, and we are in the process of appointing ours. But your Commission, unlike ours, will have a statutory duty to secure, so far as it is reasonably practicable to do so, that those holding judicial office are reflective of the community in Northern Ireland.² You see nothing inconsistent in combining this with a requirement that appointment to judicial office be only on the basis of merit. We seem to have more problems with this in England and Wales. Our Commission will only have to have regard to the need to encourage diversity in the range of people available for selection.³

For alongside the dispiriting figures I have quoted, there exists a belief, voiced by some of our most senior judges, that we have one of the best judiciaries in the world – if not *the* best. Our higher judiciary do indeed score highly on what I call the four ‘in-quotients’: my colleagues are undoubtedly incorruptible; seriously intelligent; extraordinarily industrious; and fiercely independent. Many, especially amongst their own number, will attribute this to their having had to make their way in a career at the Bar where all these qualities are now essential to success. Others may wonder how closely these perceptions are also linked to more deep-rooted British - especially English – unconscious assumptions about who is ‘the best’: well-bred, well-spoken, well-educated, white males.

Whenever I meet members of the higher judiciary from elsewhere in the world, I am struck by their own high ‘in-quotient’. Yet they, of course, do not all fit our stereotype of the white, well-spoken male. And they have usually come from a much wider range of professional backgrounds and experience. A Supreme Court bench (for example in Israel or Spain) may have members promoted from the lower judiciary, or appointed from the Chief Prosecutor’s or Attorney General’s office and other Government legal service, from the Universities, as well as from private practice. My point is that we seem to be almost the only country in the world where the only lawyers who are thought fit to serve in our highest courts are those who have excelled as advocates in those courts.

Why should it matter that our judiciary is so peculiar if it is of such high quality and even fits the traditional national stereotype of the people likely to have that quality? One answer is that times have changed: the country has

² Justice (Northern Ireland) Act 2004 s.3.

³ Constitutional Reform Act 2005 s.64(1).

changed and people's expectations have changed. Another is that judging has changed.

In the early decades of the twentieth century, many appointments to the Bench were overtly political. Yet apparently this was not hugely controversial, partly at least because the work of the judges was not seen as political. They administered the criminal law, presiding over jury trials and sentencing those found guilty. They also decided civil disputes, but in nothing like the same numbers as they do today. In the 1950s, according to Robert Stevens,

“the Chancery judges were such a disagreeable group that litigation in that Division had virtually dried up”.

The judicial work of the House of Lords was mostly devoted to ‘big money’ cases. We recently did a trawl of the cases upon which Lord Cohen, the first Jewish Law Lord, sat from 1951 to 1961. These were dominated by tax, commercial, company and planning cases, with only a smattering of cases which might affect the lives of ordinary men and even fewer that might affect the lives of ordinary women. There was only one recognisably civil liberties case and no judicial review cases at all.

The quality and quantity of our work at all levels has changed dramatically since then. Three things have contributed to the greater constitutional and ‘small p’ political significance of the work of the higher judiciary:

The first is European Union membership. In practice, this does not provoke a great deal of litigation as opposed to regulation, but the litigation it does provoke has a new quality. Some of the rights of the citizens of EU countries have to be given direct effect in member states even if they have not been legislated by the domestic Parliament, or are even incompatible with domestic legislation. A recent example is our decision⁴ that a post-operative trans person had to be treated for employment purposes as a member of the sex to which she now belonged. This was because of a decision of the European Court of Justice⁵ interpreting the Equal Treatment Directive 76/207/EEC. It had nothing to do with a later decision of the European Court of Human Rights⁶ on the right of a trans person to marry in their reassigned gender.

The second, and numerically far more important development, was the explosion in judicial review of administrative action. There were some landmark House of Lords decisions in the 1960s (such as *Padfield v Ministry of Agriculture* [1968] AC 996 and *Anisminic v Foreign Compensation Commission* [1969] 2 AC 147); but things really took off in the 1970s when a new unified procedure replaced the old prerogative writs of prohibition, certiorari and mandamus and the declaration procedure. The purpose did not change. The Court is not allowed to question the merits of an executive decision lawfully taken. Its task is to ensure that the decision is not illegal, irrational or unfairly taken. But the public increasingly saw it as a possible

⁴ *Chief constable of West Yorkshire Police v A (No 2)* [2004] UKHL 21; [2004] 2 WLR 1209.

⁵ *P v S* [1996] ICR 795.

⁶ *Goodwin v United Kingdom* (2002) 35 EHRR 447.

way to challenge decisions they did not like and an ingenious legal profession persuaded the judges to develop those concepts so that the chances of success became much greater than they had previously been.

The third is of course the Human Rights Act 1998, incorporating the European Convention into United Kingdom law. This makes several important changes. First it gives us all positive new rights against the State. Public authorities are not allowed to act in a way which is incompatible with our Convention rights. If they do the courts have new powers to put it right – unless the public authority was bound by United Kingdom statute to do what it did. Second, the courts have a new duty to read and give effect to all United Kingdom legislation so as to be compatible with the Convention rights. This can affect private individuals as well as public authorities. Third, if a statute cannot be read compatibly, the Court may not strike it down but the High Court may declare it incompatible. The Government is then bound by its international obligations to do something about it, however reluctantly. The best known example is the Belmarsh case, *A v Secretary of State for the Home Department* [2004] UKHL 56; [2005] 2 WLR 87 which led to the Prevention of Terrorism Act 2005.

These three developments have certainly made the work of the higher courts entirely different from what it was in Lord Cohen's day. But that is among the more rarefied heights of the judiciary. Other developments have made the legal system much more accessible to ordinary members of the public, and increased the range of issues that can be brought to court, many of which undoubtedly have a social and 'small p' political dimension.

For example, the post-war welfare state saw access to justice in much the same way as it saw access to other social services, as something which should be made available to those who could not afford to pay for it on similar terms to those who could. By access to justice was meant access to the legal services which would enable people to gain access to the courts. The legal aid scheme had a comparatively generous means test and its scope was progressively extended until the 1970s. In the 1970s all the talk was of 'unmet legal need'; it had yet to be acknowledged that a demand-led service access to which was controlled largely by its providers might be unsustainable. People were thus encouraged to think that they could take their cases to court and did so. Even now that everything about the legal aid scheme has been curtailed, the idea that litigation should be within the grasp of ordinary people with ordinary grievances is deeply engrained.

Another factor, not unrelated to the first, is the huge increase in family breakdown and family litigation. Divorce has been made available to all, rather than the preserve of the few. Family litigation is big business. Perversely, the public have responded with an increasing reluctance to get married at all, whether or not they have children. But they still separate and feel the need to litigate. The corresponding empowerment of women has recently been matched by the emergence of a fathers' rights movement. Recent demonstrations on behalf of separated fathers have clearly shown how sensitive and small p political this jurisdiction can be.

The access to justice theme is taken further by all the new tribunal jurisdictions which have been created; some of these are in the private sector, redressing the inequalities of bargaining power in the relationships between

landlord and tenant or employer and employee; most are in the public sector, giving the claimant or applicant an avenue of appeal against administrative decisions, for example on benefits, housing, immigration or asylum. Some of these jurisdictions are now enormous and at least as important in the lives of ordinary people as the regular courts of law. The proportion of women holding judicial office in them is higher than in the ordinary courts. But as a matter of principle they all have a route into the ordinary courts of law, usually through statutory rights of appeal, but also through judicial review.

In this changed world, where law is not the exclusive domain of commerce and the better off, it matters much more who the judiciary are and how they are perceived. All common law countries, which tend to recruit their judges from amongst their higher court advocates, tend to face the same difficulties in achieving a more reflective judiciary. But the common law country which has made the most serious attempts to do this is Canada. In the last two decades, their proportion of women judges has changed from something very like our present proportion to something over a quarter at all levels of the federal and provincial judiciary. They even have a female Chief Justice, Beverley McLachlin, something which I still find difficult to envisage here. She has given four reasons⁷ why she believes that increased representation of women on the bench has changed and will continue to change their legal system for the better:

“First, I believe that women on the Bench are necessary to the public confidence of our judicial systems. From the point of view of the public’s perception, the presence of women on the Bench makes the legal system less alien and more relevant to the female half of our population. The courts of our country should no longer be viewed as all male preserves into which females and minorities venture at their peril. The paternalistic aura which once characterised our courts is fast diminishing. Our courts have become, more than ever before in history, the people’s courts. . . . The reality is that many people, particularly women, may feel less than complete trust in a system composed exclusively or predominantly of middle-aged white men in pinstriped trousers. They will question whether such a court can reflect the various viewpoints and values of an increasingly pluralistic society. . . .”

This is closely akin to the conclusion I reached in a similar paper in 2000.⁸ In a real democracy, judges can no longer claim legitimacy solely on the basis of their legal expertise and social superiority. They have to reflect the fact that they are servants of a diverse society all of whose members are equally valued. Chief Justice McLachlin continues with a related but distinct consideration:

“The second reason I think it is important to have women on the Bench is symbolic. In a world where one of the primary functions of the judiciary is to promote equality and fairness, it

⁷ At a seminar run by the Association of Women Barristers, House of Commons, London, United Kingdom, Wednesday, July 2, 2003.

⁸ ‘Equality and the Judiciary: Why should we want more women judges?’ [2001] PL 489, delivered to the Society of Legal Scholars annual conference in 2000.

would be anomalous if the very instrument charged with that goal should itself exclude women from its ranks . . .”

Of course, it has always been the role of the judiciary to promote procedural fairness. The idea that they might also promote substantive fairness and equality is much more recent and by no means universally accepted. Chief Justice McLachlin is speaking of a legal system which has had a Charter of Fundamental Rights and Freedoms since 1982 (and statutory protection for human rights and fundamental freedoms for much longer than that). But our own legal system has been promoting social justice at least since the institution of the welfare state and is now turning its attention to fundamental human rights. Article 1 of the Universal Declaration of Human Rights declares that ‘all human beings are born free and equal in dignity and rights’. This is often demonstrably untrue in fact, but it is still an important principle and aspiration. I do believe that there is some symbolic value in the fact that our doorkeepers now shout ‘their lordships and her ladyship’ when the law lords appear on the committee corridor to file into the appellate committee room in the House of Lords.

Even those who are still having difficulty with the principle of equal dignity would probably now accept Chief Justice McLachlin’s third reason:

“The third reason for putting women on the Bench is utilitarian. Simply put, it represents a sound use of human resources. It seems to me that modern societies cannot afford to lose the intellectual power and energy of half the population. . . . Our society is increasingly complex, our birth rates are low. We need the wisdom, not only of wise men, but of our wise women.”

On the whole, the idea that all should have an equal opportunity to compete in the public arena on the basis of merit is now widely accepted. It still leaves plenty of room for argument about what we mean by merit and how it should be recognised.

Her last reason is more controversial. It is that women judges might actually make a difference:

“The fourth and most important reason why I believe we need women on our Benches is because we need the perspectives that women can bring to judging.”

Both Dr Kate Malleson⁹ and I have argued that we should not *expect* women judges to ‘make a difference’ in the sense that they are likely to make different decisions from men. Women are as diverse as men in their characters and attitudes. We are all lawyers and judges first and men or women second. We all swear the same judicial oath,

‘to do right to all manner of people, according to the laws and usages of this realm, without fear or favour, affection or ill will.’ In most cases, according to Mary Jean Coyne, Chief

⁹ ‘Justifying Gender Equality on the Bench: Why Difference Won’t Do’ (2003) 11 *Feminist Legal Studies* 1.

Justice of Minnesota, 'A wise old woman and her wise old man reach the same conclusions.'

More fundamentally, does not the idea that the gender of a judge might make a difference strike at the very idea of judicial neutrality? Many people think of the law as a collection of clear rules and principles which will lead to a determined outcome once they are applied to the proven facts of a given case. This is not in fact how the common law works, as the great Lord Reid explained in his famous lecture to the Society of Public Teachers of Law in 1971:¹⁰

"There was a time when it was thought almost indecent to suggest that judges make law – they only declare it. Those with a taste for fairy tales seem to have thought that in some Aladdin's cave there is hidden the Common Law in all its splendour and that on a judge's appointment there descends on him knowledge of the magic words Open Sesame. Bad decisions are given when the judge has muddled the password and the wrong door opens. But we do not believe in fairy tales any more."

And yet Erica Rackley¹¹ argues that people may need to think of the law as a simple set of clear rules and the judges as their neutral interpreters:

"... maybe we do need to believe in fairy tales. Maybe... a belief in the superhero judge who comes with a built in programme, a game plan to ensure a coherent and certain outcome consistent with the values and premises of the particular political tradition he is there to serve and preserve is intrinsic to our notion of judging... The merest glimmer of recognition that judges may be political actors with substantial power and opportunity to enact their personal political preferences surely threatens to render unstable the whole edifice of the law... Hence the importance of preserving the mythological dimension of the adjudicative process, ensuring its distance from the concerns of mere mortals. We can imagine the judge in no other way. He has to be seen as 'supra' human. We even make him dress up in his own kind of cape and mask – well wig – his own 'superhero' outfit."

The wig we make the judges wear is, of course, a man's wig. The very idea of a judge as a real human being with a life of her own will threaten this super-hero image. Just as Henry Cecil said that we should not be able to imagine a judge having a bath, we should not be able to imagine her doing the washing up (I wish). Rackley goes on to compare the woman judge with Andersen's little mermaid - who traded her beautiful voice for legs so that she could join her handsome prince on dry land and then found that he was no longer interested in her:

¹⁰ 'The Judge as Lawmaker' (1972) 12 *JSPTL* 22.

¹¹ In her wonderful paper on 'Representations of the (woman) judge: Hercules, the little mermaid, and the vain and naked Emperor' (2002) 22 *Legal Studies* 602, 616.

“She [the woman judge] too remains cast as a mermaid. Her physical appearance threatens to upset aesthetic norms; her presence is an inescapable irritant, simultaneously confirming and disrupting the established masculinity of the bench. As such, the woman judge is almost a contradiction in terms. She is *so* deviant that she is inevitably subject to an irrepressible desire to conform. Like Andersen’s mermaid, she is induced to deny herself and sell her voice; her dangerous siren call is silenced and in the silence difference is lost.”

This may strike a chord with those who remember our first woman High Court judge, Dame Elizabeth Lane, but it certainly would not with those who remember our second, Dame Rose Heilbron.

But although I argue that we should not *expect* a woman judge to make a difference, I do believe that having women judges can actually make a difference. This is on at least two levels, style and substance. We should not underestimate the importance of different styles of judging. They have a contribution to make to the reflective and symbolic functions of diversity mentioned earlier. We have to challenge the notion that the only person who can be taken seriously as a neutral and fair-minded person is the judicial equivalent of a tall man in a suit. We have to establish the claims of people who are recognisably women, and of people of colour and people with disabilities, to be regarded as serious and responsible. We have to reflect ourselves in the way that we speak. I speak, as I said in a recent opinion, as “a reasonable but comparatively weak and fearful grandmother” with as much right to be heard as the stronger and braver grandfathers around me.¹² We also have to stop dressing up as old men and prove that we can do the job properly without the disguise. Fortunately, in the House of Lords we have discarded the superhero outfit and dress more or less as ourselves.

Window dressing is important, as every retailer knows, but it is also desirable to improve the quality of the products on sale. We might expect a more diverse judiciary also to have some impact upon the substance of the law. As Chief Justice McLachlin put it:

“. . . jurists are human beings, and, as such, are informed and influenced by their backgrounds, communities, and experiences. For cultural, biological, social and historic reasons, women do have different experiences than men.”

Orally, she put it more graphically:

“We lead women’s lives: we have no choice.”

The experience of leading those lives should be just as much part of the background and experience which shapes the law as the experience of leading men’s lives has been for centuries. Of course, the whole idea that a judge’s experience may shape the law calls into question the role of a judge in making the law. We may not believe in the declaratory ‘fairy tale’ any more, but we certainly believe that there are limits to judicial creativity. This is a large subject upon which a great deal has been written and will be written in future. We are at the moment considering whether it is possible

¹² *R v Z* [2005] UKHL 22, [2005] 2 WLR 709, at para.73.

for the House of Lords to overrule a previous decision for the future only, rather than with retrospective effect, as is the general rule. Prospective overruling does, of course, look far more like changing the law than simply divining what the law is by the extrapolation of known rules and principles. (That is why the liberal judges of the United States Supreme Court were prepared to countenance it in the 1960s, whereas the present court is not.) In general, while we recognise that there are situations in which the courts can develop the common law to meet changing circumstances, we draw the line at judicial law reform – that should be left to Parliament. And where Parliament has already spoken, the judicial task is one of interpretation – but even here there are often choices to be made. So I want to illustrate the difference that diversity might make by two examples, one of a development in the common law, and the other of a development in statutory interpretation.

The great majority of judgments which I have written or spoken could just as easily have been written or spoken by a man. But there is one exception. In *Parkinson v St James and Seacroft University Hospital NHS Trust* [2001] EWCA Civ 530, [2002] QB 266, I tried to put into words the experience of bearing and rearing a child from the woman's point of view. *Parkinson* was an example of a knotty problem in an almost completely judge-made area of the law: what should be the scope of a doctor's liability if she negligently fails to sterilise a man or a woman properly so that a child is conceived and born whom the parents never intended to have and had good reasons for not wishing to have? Should the doctor (or rather the doctor's employers or insurers) have to pay the full cost of bringing that child up? Or should the parents have to bring into account the incalculable benefit which such a child can bring to them and to society? And what if the child is disabled, so that he costs more to bring up than other children do? And what if the mother is disabled, so that she is unable to bring up the child without a great deal of extra help?

After initial wobbles, the English Court of Appeal decided, applying entirely orthodox principles, that the doctor had to pay. She undoubtedly had a duty to take reasonable care to perform an effective sterilisation or warn that it might not be so. The entirely foreseeable result of failing to take such care was an unwanted pregnancy. The pregnancy brought with it identifiable consequences which would ordinarily sound in damages. The benefits could not be quantified so could not be taken into account. The House of Lords, in *MacFarlane v Tayside Health Board* [2000] 2 AC 59, eventually took a different view. By then they had developed a new doctrine to limit the scope of liability in negligence. They had done this largely because they had extended the scope of the tort to cover financial loss as well as physical harm to persons or property. The duty to take care now extends only so far as it is fair, just and reasonable to impose it.¹³ Hence the House held that it was fair, just and reasonable to compensate the mother for the pain and suffering involved in the pregnancy and childbirth and for the immediate costs associated with this. But it was not fair, just and reasonable to compensate the parents for the full cost of bringing up the child they had never meant to have. In *Parkinson*, in the Court of Appeal, we held that the parents could

¹³ See *Caparo Industries plc v Dickman* [1990] 2 AC 605.

nevertheless have the extra costs involved in bringing up a disabled child. And in *Rees v Darlington Memorial Hospital* [2003] QB 20 we held that a blind mother could have the extra costs involved in bringing up the child she had sought to avoid having precisely because she did not feel able to bring one up herself. Three of a seven judge House of Lords¹⁴ held that we were right: the *McFarlane* decision had only been concerned with the upbringing of a healthy child by healthy parents. Four of them held that we were wrong: but they nevertheless acknowledged that negligently to cause a woman to conceive and bear a child she never meant to have was an invasion of her bodily autonomy which should be marked by a conventional sum in compensation, which they put at £15,000: not nominal but not the full cost or even the extra cost of bringing up the child. Where this leaves the *Parkinson* decision is not entirely clear.

Although in one sense I lost the battle, there is another sense in which I believe that I won the war. The tendency had been to conceptualise the damage caused by having an unwanted child in financial terms – the financial costs of feeding, housing, clothing and educating her. I tried to reframe the damage in terms, first of the invasion of the mother’s right to bodily integrity and autonomy, and second in terms of the responsibility which having any child entails 24/7 until the child is old enough to fend for herself.

Even when it comes to statutory interpretation there are choices to be made, in which the different perspectives of the different judges may also be helpful. A good example is the recent fate of paragraph 2 of the Rent Act 1977 as amended by the Housing Act 1988. This gives preferential rights to succeed to a Rent Act tenancy to the surviving spouse of a deceased tenant and also to a person who was “living with the deceased tenant as his or her wife or husband” at the date of death. It gives less preferential rights to a “member of the deceased tenant’s family” who was living there at the date of death. So how should the law treat the surviving partner in a homosexual relationship?

There is no doubt that when the concept of a member of the deceased tenant’s family was first enacted in 1920 Parliament could not possibly have intended that it would cover the surviving partner in a homosexual relationship. Sodomy and gross indecency were still serious criminal offences. But in *Fitzpatrick v Stirling Housing Association* [2001] 1 AC 27, a majority of the House of Lords held that it now did so. They employed the principle that a statute is ‘always speaking’, so that a word which would not have applied to this situation when it was first enacted had to be interpreted in the light of contemporary social circumstances. This in itself was quite a bold decision, but their lordships stopped short of holding that the survivor had been living with the tenant as his or her wife or husband.

Yet in *Ghaidan v Godin-Mendoza* [2004] UKHL 30; [2004] 2 AC 337, we felt able to take that further step. We found that Parliament had now told us to do so by enacting the Human Rights Act. Article 8 of the European Convention guarantees the right to respect for a person’s home. Article 14 requires that the enjoyment the Convention rights and freedoms be secured

¹⁴ [2003] UKHL 52; [2004] 1 AC 309.

without discrimination on grounds such as race, sex, etc. This does not mean that everyone is entitled to a home. But it does mean that if the State steps in to protect the home, it must not discriminate unjustifiably in the classes of people to whom it gives such protection. We could see that there might be a good reason to discriminate between married and unmarried couples, but we could not see any relevant difference between unmarried heterosexual and unmarried homosexual couples for this purpose. Nor could we see any justification for treating one less favourably than the other. We now have a duty to interpret legislation, if possible, so as to be compatible with the Convention rights even if this is not accordance with what Parliament originally intended. Was it possible to interpret the words 'as his or her wife or husband' to include a homosexual partner? One of us thought not: the words 'wife' and 'husband' necessarily connote having a partner of the opposite sex. The rest of us thought that it was: the law no longer ascribes defined roles to husband and wife, so we would be spared the problem of working out which was which, and could simply look for the necessary qualities of permanence, commitment and inter-dependence which make a relationship marriage like.

I am not suggesting that my sex had much to do with this. But an awareness of gender and discrimination issues may have had something to do with it, as may a background in family law and the experience of sitting as a family judge. When I speak of diversity on the bench, I am not only speaking of diversity in the obvious characteristics of sex and ethnicity, but also of diversity in legal experience and expertise.

The two examples I have given show how both the Court of Appeal and House of Lords can play an important part in making and shaping the law. They are both collegiate bodies. The Court of Appeal sits usually in threes, the House of Lords in a minimum of five, and sometimes seven or nine. Efforts are made to secure a suitable balance of expertise in the Court of Appeal. Efforts are made to secure an appropriate ethnic balance in the House of Lords, in the sense of appropriate representation from all three constituent parts of the United Kingdom. But otherwise few efforts seem to be made to ensure an appropriate balance of expertise in the House of Lords: otherwise it would have been taken for granted that there should be crime, public law and family law specialists there at all times.

In my view the new Supreme Court, like other Supreme Courts the world over, should aspire to a greater mix of gender, ethnicity and legal experience and expertise. None of this is incompatible with appointment on merit: that has always depended upon how you define and recognise merit. The Canadians found that in order to secure a more diverse judiciary they had to confront the common law's traditional definition of merit – outstanding ability as an advocate at the highest level. That traditional definition is closely related to the ease with which those who have had traditionally had a dominant role in the selection process can identify the people who have it. In other words, 'we all know who the good people are'. They are all, of course, excellent people. But I share Lord Falconer's view that I cannot possibly be the first woman lawyer who has been (if I am) good enough to join them. I was merely the first who was visible to them.

So if we are to take serious steps towards a more diverse judiciary, we need to do three things. The first is to redefine 'merit': we have to look for the

best and brightest lawyers, with the right mix of skills and qualities, from all walks of legal life, not just from the ranks of the higher advocates or litigators. The second is therefore to take positive steps to spread the net wider and to encourage more people to make themselves visible to the appointing bodies. And the third is to introduce proper procedures for making those appointments: procedures that would be recognised in the professional world outside the law as reflecting the best equal opportunities practice. But we do also need to retain the flexibility to be a little bold.

The arguments for this sort of diversity, especially at the highest law-making judicial level, are not unlike the arguments for selecting juries from a fair cross section of the community. The United States Supreme Court made much the same point in *Ballard v US* (1946) 329 US 187, when deciding that a jury panel composed predominantly of men was not a fair cross section of the community:

“The thought is that the factors which tend to influence the action of women are the same as those which influence the action of men – personality, background, economic status – and not sex. Yet it is not enough to say that women when sitting as jurors neither act nor tend to act as a class. Men likewise do not act as a class. But if the shoe were truly on the other foot, who would claim that a jury was truly representative of the community if all men were intentionally and systematically excluded from the panel? *The truth is that the two sexes are not fungible*; a community made up exclusively of one is different from a community composed of both; the subtle interplay of influence one on the other is among the imponderables. To insulate the courtroom from either may or may not in a given cause make an iota of difference. Yet a flavour, a distinct quality is lost if either sex is excluded. The exclusion of one may indeed make the jury less representative of the community than would be true if an economic or racial group were excluded.”

A similar point was made by Benjamin Cardozo in his seminal lectures on judging first published in 1921:¹⁵

“The eccentricities of judges balance one another. One judge looks at problems from the point of view of history, another from that of philosophy, another from that of social utility, one is a formalist, another a latitudinarian, one is timorous of change, another dissatisfied with the present; out of the attrition of diverse minds there is beaten something which has a constancy and uniformity and average value greater than its component elements.”

So I have absolutely no intention of turning into the little mermaid. I have every intention of doing my bit to contribute to the attrition of diverse minds.

¹⁵ See *The Nature of the Judicial Process*, Yale University Press, 1961, p.177.

THE CONFIDENTIALITY OF MEDICAL RECORDS: INFORMATIONAL AUTONOMY, PATIENT PRIVACY, AND THE LAW

Joan Loughrey, Lecturer, School of Law, University of Leeds

INTRODUCTION

This article examines critically the level of protection currently afforded to medical privacy by the law of confidence in situations involving the use and handling of personal information by NHS bodies. Patient data is used by the NHS for a broad range of purposes, not limited to care of the patient, including administrative, auditing and educational purposes.¹

A number of developments in recent years have combined to threaten patient privacy. The increasing utilisation of information technology is making the collation and dissemination of patient data easier both within and without the NHS. Electronic Care Records are being introduced which will contain a lifelong record of a patient's health and healthcare.² Just as locating patient information becomes easier, the nature of that information, and its sensitivity, is altering as developments in the science of genetics impacts on health care.³ Furthermore recent organizational reforms in the NHS may undermine medical privacy.⁴ There have been various legislative initiatives aimed at removing legal obstacles to the transmission of patient data.⁵ New Labour's drive for better integrated public services has required greater sharing of information between public agencies for a wide range of administrative purposes.⁶ The Government has indicated that the data collected under improved information systems will facilitate NHS management processes, will be made available to local authority social

¹ See Department of Health, *Confidentiality NHS Code of Practice* (November 2003) Annex C, which sets out a wide range of situations in which patient data may be disclosed by the NHS.

² It is envisaged that, at some point, the Electronic Patient Record may include the results of tests carried out upon patients pre-birth: Human Genetics Commission, *Inside Information, Balancing Interests in the Use of Personal Genetic Information* (May 2002) Ch.4, para.4.21.

³ For a detailed review of the legal and ethical aspects of privacy in the context of genetic information see G.T. Laurie, *Genetic Privacy: a Challenge to Medico-Legal Norms* (2002).

⁴ Perri 6, C. Raab and C. Bellamy, "Joined-Up Government and Privacy in the United Kingdom: Managing Tensions Between Data Protection and Social Policy: Part I" (2005) 83 *Public Administration* 111 at p.114.

⁵ Notably, the Health and Social Care Act 2000, s.60, the Health Service (Control of Patient Information) Regulations 2002, SI 2002/1438 and the Health and Social Care (Community Health and Standards) Act 2003. For a critique of the first two, which are intended to be transitional measures, see P. Case, "Confidence Matters: The Rise and Fall of Informational Autonomy in Medical Law" (2003) 11 *Medical Law Review* 208.

⁶ Perri 6, Raab and Bellamy, n.4 above, at pp 112-113.

services, and will further constitute an exploitable commercial asset.⁷ At the same time, given demographic changes and the growing significance of health services in the lives of many citizens, the ability of patients to control medical information is arguably an increasingly important aspect of individual autonomy.

These developments are problematic for a number of reasons. The empirical evidence in Britain suggests that both communities and patients place a high value on medical privacy.⁸ While most people have no objection to the disclosure of information within the NHS for the purpose of their own treatment, there is general unease about the sharing of such data more widely with other healthcare professionals, social workers, NHS managers or support staff.⁹ The Government's own research shows that individuals do not feel positively about the disclosure of medical information to achieve claimed indirect benefits such as better functioning of public services, or for purposes such as audit.¹⁰ Furthermore, a majority of individuals wish to have the opportunity to consent to the various uses made of their healthcare data, and believe that refusals of consent should be respected.¹¹ The current law in this area may therefore be regarded as deficient to the extent that it fails to give effect to public expectations about respect for informational autonomy and control of access to health records.

The principle that an individual's autonomy should be respected is fundamental to medical law and ethics. It requires that an individual is able to make decisions about their life free from coercion or impediment and that such decisions are respected. Informational autonomy constitutes the ability

⁷ Department of Health, *Building the Information Core-Implementing the NHS Plan* (January 2001) para.4.5, para.5.6. and para.6.14 respectively. See also Department of Health, *Confidentiality NHS Code of Practice* (November 2003) Annex C. For a discussion of the growing pressures on informational privacy see R. Pattenden, *The Law of Professional-Client Confidentiality* (2003) Ch.2; Perri 6, Raab and Bellamy, n.4 above.

⁸ European Opinion Research Group (Special Eurobarometer), *Data Protection* (December 2003) at p.7: 75 % of the British public were fairly or very concerned that their privacy should be protected.

⁹ For a comprehensive survey of studies on the public attitude to medical privacy see NHS Information Authority, *ERDIP Evaluation Project: N5-Patient Consent and Confidentiality Study Report* (16th May 2002): one study, involving 1239 responses, found that 40% objected to disclosure to hospital managers, social workers and practice support staff: at 29; NHS Information Authority, *Share with Care! People's views on Consent and Confidentiality of Patient Information* (October 2002) at pp.9, 13-14, 26: only 17% were happy for NHS managers to access records.

¹⁰ Performance and Innovation Unit, *Privacy and Data Sharing: the Way Forward for Public Services*, Annex C para. C27 (hereafter "PIU"), although the report noted 1999 research by the National Consumer Council which had concluded that there was little public concern about patient records circulating within the NHS: at para.C11.

¹¹ NHS Information Authority, n.9 above, at pp.9, 13-14, 26. See also M. Robling, K. Hood, H. Houston, R. Pill, J. Fay and H. Evans, "Public attitudes towards the use of primary care patient record data in medical research without consent: a qualitative study" (2004) 30 *Journal of Medical Ethics* 104; PIU, n.10 above, at para.C23.

“to control what is known (and by whom) regarding oneself and one’s activities.”¹² It has been argued that this “informational self determination” constitutes the core interest protected by the law of confidence.¹³ When an autonomous person’s informational privacy has been breached, that is, where her personal private information has been disclosed without her consent, there is a failure to respect her autonomy.

A failure to respect autonomy is morally objectionable because it fails to respect an individual as a person and as a rational moral agent¹⁴ and uses her as a means to an end rather than as an end in herself.¹⁵ Consequently, insofar as the law permits the disclosure of health care records without giving due weight to the principle of informational autonomy, it may be regarded as morally deficient.

Against this background the paper argues that the law fails adequately to protect individual interests in medical privacy and that, contrary to expectations generated by the Human Rights Act 1998, there is little difference between the protection of informational autonomy under Article 8 jurisprudence and that provided by the common law. Despite some indications that the courts are beginning to recognise the dignitary interests underpinning the law of confidence, this is unlikely to result in increased protection of patient privacy and informational autonomy due to the reluctance of the judiciary to intervene in administrative decision-making where disclosure occurs or is sought by public bodies ‘in the public interest’.

The main focus of the paper is on the common law of confidence in the context of changes brought about under the Human Rights Act 1998 (“the HRA”). Although the Data Protection Act 1998 (“DPA”) may be interpreted to give effect to Article 8 rights to informational privacy¹⁶ and although it introduced significant safeguards directed at the manner in which public bodies deal with data, it may have little to add to the common law in the extent to which it advances informational autonomy. While it is true that it grants patients additional rights to control their data including a right to prevent its use,¹⁷ these rights will be of little practical use in the medical context since there remains a lack of patient awareness of the range of

¹² J.R. Pennock, “Introduction” in J.R. Pennock and J.W. Chapman (eds), *Privacy, NOMOS XIII* (1971), p.xiii.

¹³ G. Phillipson and H. Fenwick, “Breach of Confidence as a Privacy Remedy in the Human Rights Act Era” (2000) 63 *MLR* 660 at p.662. It should be noted that informational autonomy and privacy are not the same. For example privacy may be protected in circumstances where individuals have been given no rights to control information access and, it will be seen, protected for reasons other than concerns with informational autonomy. However a key aspect of recognising informational autonomy is that it not only provides a reason for respecting the privacy of the autonomous, it requires that control over information is exercised by the persons to whom it relates.

¹⁴ S. Benn, “Privacy, Freedom and Respect for Persons” in Pennock and Chapman, n.12 above, Ch.1; J.S. Mill, *On Liberty* (1869) Ch.I and Ch.III.

¹⁵ I. Kant, *Foundations of the Metaphysics of Morals*, (1959) (trans. Lewis White Beck).

¹⁶ As a result of the courts’ obligation to interpret legislation compatibly with the Convention wherever possible: Human Rights Act 1998, s.3.

¹⁷ Data Protection Act 1998, s.10.

disclosures which may take place.¹⁸ Patients are not, therefore, in a position to exercise their rights under the DPA to prevent disclosure. Should they discover subsequently that disclosure has occurred, only the law of confidence provides them with a remedy. Under the DPA, once information is released, compensation for distress alone will usually be unavailable, a restriction which casts doubt on its efficacy as a remedy for invasions of privacy.¹⁹ Furthermore the DPA contains a range of defences and exceptions which legitimise the use of data without patient consent and even in the face of patient objections, which undermine the extent to which it can be relied upon to advance informational autonomy.²⁰ However in all cases processing must be lawful, that is, it must not be carried out in breach of an obligation of confidence.²¹ Consequently the law of confidence occupies a central role in the protection of privacy.

The article is structured as follows: firstly the rationales for the protection of medical confidentiality both before and after the Human Rights Act 1998 will be assessed including the extent to which the value of informational autonomy has been recognised by the law and whether such recognition is necessitated by Article 8 of the HRA. Secondly the scope of the public interest defence and Article 8(2) justifications for disclosure will be examined to ascertain whether the need to justify privacy interferences by reference to Article 8(2) affects the protection afforded to privacy. Finally, to the extent that informational autonomy is acknowledged, the paper will analyse why it will not result in better protection for privacy where access to information is sought by the State.

MEDICAL CONFIDENTIALITY AND THE PROTECTION OF AUTONOMY

Pre- Human Rights Act 1998: the Public Interest Rationale

Briefly the traditional formulation of the action for breach of confidence was, “the information . . . must have the necessary quality of confidence about it. Secondly, that information must have been imparted in circumstances importing an obligation of confidence. Thirdly, there must be an

¹⁸ NHS Information Authority, n.9 above.

¹⁹ Data Protection Act 1998, s.13. There are exceptions where the data has been processed for journalism, artistic purposes or literary purposes: Data Protection Act 1998, s.3; See criticism in *Martin v United Kingdom* (2003) 37 E.H.R.R. C.D. 91, at 100.

²⁰ For example, a person can only prevent processing if they are able to show that it will cause substantial and *unwarranted* distress: Data Protection Act 1998 s.10. For further discussion of the operation of the DPA in the medical context see M. Brazier, *Medicine Patients and the Law* (2003), pp. 76-77; A. Grubb, *Principles of Medical Law* (2004), paras.9.192-9.193. See also L. Hagger, S. Woods and P. Barrow, “Autonomy and Audit-Striking the Balance” (2004) 6 *Medical Law International* 105 which demonstrates how the DPA can be interpreted to permit the use of data without consent.

²¹ Data Protection Act 1998 Sch.1. Information Commissioner, *Use and Disclosure of Health Data: Guidance on the Application of the Data Protection Act 1998* (May 2002) at Ch.4.

unauthorised use of that information to the detriment of the party communicating it.”²²

In the context of medical information, the doctor/patient relationship is one of trust and has long been protected by the law of confidence.²³ In relation to other health care professionals, the therapeutic context has been held to be sufficient to import an obligation of confidence, whether or not health care professionals have treated the patient.²⁴

Historically, the public interest has dominated judicial explanations for protecting confidentiality.²⁵ Examples include *X v Y* in which a newspaper sought to publish the identities of two GPs who had been diagnosed with AIDS.²⁶ The information had probably been disclosed by a hospital employee in breach of a contractual obligation of confidence. The hospital was granted an injunction on the basis that there was a public interest in protecting confidentiality which comprised encouraging persons to come forward for treatment and the consequent protection of public health.²⁷

Subsequently, in *W v Egdell*, an expert medical report concerning a dangerous prisoner was disclosed in breach of confidence. The Court of Appeal rejected the argument, accepted at first instance, that the right to confidence was based upon an individual’s private interest. The Court, echoing the reasoning in *X v Y*, affirmed that the maintenance of confidentiality was based upon the public interest.²⁸

Given the nature of this rationale, the common law naturally developed a defence to an action for breach of confidence where it could be shown that disclosure was, overall, in the public interest.²⁹ The scope of the public interest defence is unclear and, prior to the HRA, it had only been judicially recognised as operating to justify disclosures to protect the public from harm, to prevent or detect a crime,³⁰ or to disclose an iniquity,³¹ although in the context of medical confidentiality, the precise scope of the last two categories remains uncertain. It is unclear, for example, whether the public interest defence is confined to serious crime. Professional guidance leaves

²² *Coco v A.N. Clark (Engineers) Ltd* [1969] R.P.C. 41, at 47. This paper will not analyse in detail the elements of the action in confidence nor its subsequent development. This has been dealt with elsewhere. See Grubb, n.20 above, at Ch.9.

²³ *Hunter v Mann* [1974] 1 Q.B. 767, at 772; *X v Y* [1988] 2 All E.R. 648; *W v Egdell* [1990] Ch.359; *AG v Guardian Newspapers (No. 2)* [1990] 1 A.C. 109 H.L., at 255 (Lord Keith).

²⁴ *Re C (a minor) (wardship: medical treatment)* [1989] 2 All E.R. 791, at 795.

²⁵ *Attorney General v Guardian Newspapers Ltd (No 2)* n.20 above, at 256 (Lord Keith), at 269 (Lord Griffiths), at 281 (Lord Goff).

²⁶ n. 20 above.

²⁷ There was also a public interest in encouraging loyalty in health service employees to their employer.

²⁸ n. 20 above, at 416 (Sir Stephen Brown P.), at 420 (Bingham L.J.).

²⁹ *Attorney General v Guardian Newspapers (No 2)*, n.20 above, at 269 (Lord Griffiths), at 282 (Lord Goff); *X v Y*, n.20 above; *W v Egdell*, n.20 above.

³⁰ *W v Egdell*, n.20 above, at 425.

³¹ *Beloff v Pressdram Ltd* [1973] 1 All E.R. 241; *Initial Services v Putterill* [1968] Q.B. 396.

the matter to the discretion of the health care professionals and the case law has not addressed the point.³²

It was also unclear whether uses of patient data for NHS purposes, such as teaching, financial and clinical audit, management and administrative purposes and medical research could be justified on the basis of this defence.³³ There had been *dicta* outside the sphere of medical confidentiality which suggested that a broad public interest defence of just cause or excuse existed which could legitimate disclosures where the public interest in disclosure outweighed the interest in maintaining confidence, even though no wrongdoing is involved.³⁴ If correct, these comments could have justified a broad range of disclosures.

In the end, the legality of such uses was not tested. This was not because such disclosures did not occur but rather because patients did not litigate over them. However, had a challenge to such uses been mounted, it is likely to have been unsuccessful. There are problems in seeking to protect rights by appealing to public interest justifications. Lyons argues that basing a legal right upon utilitarian grounds cannot accommodate the right having moral force. The presumptive moral force of a right means that direct utilitarian arguments, which would otherwise justify conduct, must be excluded where the conduct interferes with the right. However where a right is justified on the basis that it promotes utility it can have no presumptive force, as considerations of direct utility will always be relevant in determining whether the right should be respected or not.³⁵ It might be

³² Although, in *Campbell v M.G.N.*, the defendant relied on Ms Campbell's presumptive criminality in possessing and using Class A drugs to justify publication, the point was not addressed by the court: [2003] EWCA Civ 1373; [2003] Q.B. 633 at 676. In the only other medical case which raises the point, *W v Egdell*, n.20 above, there was a real risk to public safety should the types of offence anticipated have occurred. The Department of Health, n.1 above, at pp.34-35, advises that disclosure can occur for the detection, investigation and punishment of serious crime and/or to prevent serious harm, which includes crimes involving substantial financial gain or loss. The guidance also indicates that disclosure may take place where there would be a significant benefit to the public. See also GMC, *Confidentiality: Protecting and Providing Information* (April 2004), at paras. 22-27.

³³ In *R. v Department of Health, ex p. Source Informatics Ltd* [2001] Q.B. 424, at 444, Simon Brown L.J. considered, albeit reluctantly, that the defence could cover uses for management purposes. See also *A Health Authority v X* [2001] 2 F.L.R. 673, at 696; I. Kennedy and A. Grubb, *Medical Law* (2002), p.1118; GMC, *Research: The Role and Responsibilities of Doctors* (2002), at paras. 37-39.

³⁴ *Fraser v Evans* [1969] 1 All E.R. 8, at 10; *Woodward v Hutchins* [1977] 2 All E.R. 751, at 753 (Lord Denning); *Lion Laboratories Ltd v Evans* [1985] Q.B. 526, at 536 (Stephenson L.J.); *D v NSPCC* [1978] A.C. 171, at 230 (Lord Hailsham); *Attorney General v Guardian Newspapers Ltd (No. 2)*, n.23 above, at 282 (Lord Goff). For a review of the case-law see Pattenden, n.7 above, at paras. 11.41-11.52. Both Wacks and Gurry argue that the cases in which such a suggestion has been made are of weak authority: the defences of disclosure of an iniquity or prevention of harm were also available, or there existed some element of wrongdoing or risk to the public: R. Wacks, *Personal Information, Privacy and the Law* (1989), pp.115-117; F. Gurry, *Breach of Confidence* (1984), pp.334-341.

³⁵ D. Lyons, "Utility and Rights" in J. Waldron (ed.), *Theories of Rights* (1984) p.110, at p.113 and generally.

argued that the legal right to confidence would prevail over those considerations of general utility which do not fall within the scope of the public interest defence but it would be vulnerable to arguments that it should be set aside where the public interest would be better served by doing so. It would therefore be logical for the courts to give the public interest defence a broad interpretation. In addition where the considerations of general welfare do fall within the parameters of the public interest, then a right to confidence founded on the public interest can never provide a reason for overriding these considerations if general welfare, in the form of the public interest, is, in fact, advanced overall by disclosure. The strength of such a right depends upon the scope and weight of the public interest defence rather than upon the right itself. Furthermore this approach is vulnerable to the common criticisms levelled at utilitarianism. As it is not possible to objectively measure one outcome against another in order to ascertain which maximises overall welfare, the protection afforded to confidential information on this basis is likely to be subjective, *ad hoc* and uncertain.

In summary, prior to the Human Rights Act 1998, informational autonomy was poorly respected as a matter of practice in the NHS. Furthermore, it was not a value promoted by the law. Moreover the strength of the legal right to medical confidentiality, that is, its ability to withstand arguments in favour of disclosure in order to benefit the public, though untested, was probably weak.

After the Human Rights Act 1998: The Three Rationales

Following the HRA, the law of confidence has developed as the main vehicle for protecting the Article 8 right to a private life.³⁶ This has resulted in the strengthening of the action for breach of confidence in a number of respects.³⁷ For example, the obligation of confidence can now arise independently of any relationship of confidence, where the recipient of the information knows or ought reasonably to know that the information is confidential.³⁸ Where an unjustified breach is established damages for mental distress are now available.³⁹ However the action for breach of confidence continues to provide only imperfect protection for informational privacy.⁴⁰ The obtaining of information by the NHS (for example, where it is

³⁶ *Campbell v M.G.N.* [2004] UKHL 22; [2004] 2 A.C. 457, at 465 (Lord Nicholls), at 472-473 (Lord Hoffmann), at 486 (Lord Hope); *Douglas v Hello* [2005] EWCA Civ 595, at para.53.

³⁷ Although, conversely, the insistence that privacy rights must be shoehorned into the action for breach of confidence has weakened the potential scope of Article 8 and the protection of privacy: *Peck v United Kingdom* (2003) 36 E.H.R.R. 41, at paras. 111-113;

³⁸ *Campbell v M.G.N.*, n.36 above.

³⁹ *Cornelius v De Taranto* [2002] E.M.L.R. 6; *Douglas v Hello! Ltd (No 3)* [2003] EWHC 786; [2003] 3 All E.R. 996, affirmed on appeal [2005] EWCA Civ 595; [2005] 2 F.C.R. 487. *Campbell v M.G.N.*, n.36 above. For an overview of developments in this area see Pattenden, n.7 above, at paras. 8.55-8.65 and Update <http://www.uea.ac.uk/law/resources/8-28.htm> (last visited 24 April 2005).

⁴⁰ And wholly inadequate for protecting other forms of privacy such as freedom from intrusion and spatial privacy: see *Wainwright v Home Office* [2003] UKHL 53; [2004] 2 A.C. 406, *Martin v UK* (2003) 37 E.H.R.R. C.D. 91 at 100 and Laurie, n.2 above.

obtained as the result of clinical observation) and its retention, rather than use or disclosure may not be a breach of confidence for example, nor is legislatively authorised disclosure, although this may be challenged as an interference with Article 8 rights.⁴¹

It has been argued that the incorporation of Article 8 requires greater protection of the value of informational autonomy. This view stems from equating the protection of informational privacy with the protection of informational autonomy and the right to control the release of personal information.⁴²

It does not follow however that because the law is now concerned with protecting informational privacy, albeit imperfectly, it must also be concerned with the protection of informational autonomy. There is little consensus on what privacy is, or why it should be protected, and competing explanations of its rationale are available which could equally well underpin the law.⁴³ Some consider that privacy is worthy of protection on utilitarian grounds, that invasions of privacy may hinder the good life, for example, by hindering the pursuit of desirable private activities which produce social goods.⁴⁴ Others justify the protection of privacy because it is instrumental in protecting other values such as dignity and respect for persons.⁴⁵ It can be argued, for example, that the protection of privacy is instrumental in protecting mental integrity.

In fact, while informational autonomy has received some recognition from the judiciary, it competes as a rationale with two others. The utilitarian rationale continues to appear in the form of public interest arguments in support of the obligation of confidence while a third rationale has developed in the form of the protection of mental and physical integrity. The degree to which these rationales feature in the case law, and the extent to which their

⁴¹ In *R. (Marper) v Chief Constable of South Yorkshire* [2004] UKHL 39; [2004] 1 W.L.R. 2196, the House of Lords considered that the retention of samples and fingerprints did not even infringe Article 8 but in *Chare (nee Jullien) v France* (1991) 71 D.R. 141, the European Commission held that the retention of a medical file did. See also J. Morgan, "Privacy, Confidence and Horizontal Effect: "Hello" trouble" (2003) 62 *CL.J.* 444.

⁴² See for example, A. Westin, *Privacy and Freedom* (1967) 7; E.L. Beardsley, "Privacy, Autonomy and Selective Disclosure" and H. Gross, "Privacy and Autonomy" in Pennock and Chapman (eds.), n.12 above, Ch. 3 and Ch. 9; L. Lusky, "Invasion of Privacy: a Classification of Concepts" (1972) 72 *Columbia Law Review* 693; D. Feldman, "Secrecy, Dignity, or Autonomy? Views of Privacy as a Civil Liberty" (1994) *Current Legal Problems* 41; R. Singh, "Privacy and the Media: the Impact of the Human Rights Bill" in B. Markesinis (ed.), *Protecting Privacy* (1999) p.169; Phillipson and Fenwick, n.13 above, at p.674.

⁴³ See overview R. Wacks, *The Protection of Privacy* (1980), Ch. 1 and Laurie, n.2 above, Ch. 2.

⁴⁴ M. Weinstein reviews the range of arguments on this point in "The Uses of Privacy in the Good Life" in Pennock and Chapman (eds.), n.12 above, Ch. 5.

⁴⁵ B. Neill, "Privacy: A Challenge for the Next Century" in Markesinis, n.42 above, at pp.22, 28; E.J. Bloustein, "Privacy as an Aspect of Human Dignity: An Answer to Dean Prosser" [1964] 39 *New York University Law Review* 962; Feldman, n.42 above, at p.58, who also links it with the protection of autonomy; S. Benn, "Privacy, Freedom and Respect for Persons" in Pennock and Chapman, n.12 above, Ch. 1; Laurie, n.2 above, at pp.214, 248.

recognition is compatible with the approach of the Strasbourg court to medical privacy, will now be examined.

(a) *The public interest*

In *Ashworth Hospital Authority v MGN Ltd.*, Lord Phillips stated that:

“The disclosure of confidential medical records to the press is misconduct that is not merely of concern to the individual establishment in which it occurs. It is an attack on an area of confidentiality which should be safeguarded in any democratic society. The protection of patient information is of vital concern to the National Health Service and, I suspect, to health services throughout Europe.”⁴⁶

In the House of Lords, Lord Woolf also recognised the community interest in protecting medical confidentiality so as to encourage persons to come forward for treatment.⁴⁷

The problems of basing a right to confidence upon the public interest have been explored. However a feature of case-law post the HRA has been the protection of medical confidentiality against press disclosures, including in those cases where a public interest rationale has been cited.⁴⁸ A conventional explanation of these cases is that medical confidentiality receives special protection by the courts. Given that the courts consider that freedom of the press is a powerful value to be protected and upheld, if they uphold medical confidentiality on public interest grounds despite such a weighty conflicting public interest, it might be argued that, in fact, this demonstrates that the public interest in confidentiality provides a strong basis for the protection of patient privacy.⁴⁹

However, when the case law on medical confidences is viewed as a whole, it is apparent that it is the nature of the disclosure, rather than the nature of the information alone, which is significant. In non-media cases, medical confidentiality is poorly protected. The explanation for this lies in considerations of utility. If medical confidences are protected in order to encourage people to come forward for treatment without fear of disclosure of their records, it follows that where a particular disclosure will not affect public trust, utility may not require adherence to a rule of confidence.⁵⁰ Utility does not dictate disclosure in such circumstances, nor does the rule imposing an obligation of confidence disappear. Clearly, the courts have concluded that the public interest is served by having a general rule upholding medical confidentiality, although it is subject to exceptions.

⁴⁶ [2001] 1 W.L.R. 515, at [99].

⁴⁷ [2002] 1 W.L.R. 2033, at 2051-2052, citing *Z v Finland* (1998) 25 E.H.R.R. 371 at para.95. See also *X v A Health Authority* [2002] 2 All E.R. 780, at 784; *H (a Healthcare Worker) v Associated Newspapers Ltd* [2002] E.M.L.R. 23 at para.27.

⁴⁸ *Ashworth Hospital Authority v M.G.N. Ltd.*, n.47 above; *Campbell v M.G.N.*, n.36 above, at 474 (Lord Hope), at 487 (Baroness Hale).

⁴⁹ For strong statements of the importance of press freedom see *R v Central Independent Television plc* [1994] Fam. 192, at 203; *A v B Plc* [2002] EWCA Civ 337; [2003] Q.B. 195, at 205.

⁵⁰ See H. Lesser and Z. Pickup, “Law, Ethics and Confidentiality” (1990) 17 *Journal of Law and Society* 17.

Where a rule exists, reasons must be advanced if it is to be set aside in a particular case. If no satisfactory reasons are established, the rule will be upheld.⁵¹ However, provided that some good reason for disclosure can be demonstrated, the rule will be easier to set aside when it is operating in this default mode than when the circumstances of a particular case lend it some content.

The courts have not considered that the public interest in medical confidentiality is affected by limited disclosures to public authorities and have dismissed the argument that disclosure should not occur because it could deter the individual(s) to whom the information relates from seeking further treatment.⁵² In *Re C (A Minor) (Evidence: Confidential Information)*, in permitting the use of confidential medical information in adoption proceedings, Boothman J. commented, "If I allow this affidavit in evidence, it is not going to dent public confidence in the medical profession. It is not going to result in patients having less confidence in their doctors. The mother, no doubt, will have less confidence, but the public at large certainly won't."⁵³ As the courts have denied that the interest underpinning the right to confidence has been affected in such cases, the rule has been easy to set aside. Similarly in *Z v Finland* the European Court's finding that disclosure for the purposes of the investigation would not impact on the public interest in medical privacy influenced the conclusion that the disclosures were legitimate.⁵⁴

Conversely press publication of medical information is a highly public breach of confidentiality. A failure to penalise and deter such conduct might seriously undermine patient confidence in the protection of medical information and so damage the public interest which the law is concerned to protect. In such cases the rule requiring that confidentiality should be respected would be operating in substantive, rather than default, mode and would therefore be more difficult to set aside.

Again, where disclosure to bodies other than the press has been authorised, the courts have emphasised that disclosure was made to an appropriate entity and that the material would remain confidential in the hands of the recipient.⁵⁵ Dissemination was unlikely to attract much public attention and so the public interest in confidentiality was unlikely to be affected.

⁵¹ There is a view that utilitarianism, even act utilitarianism, can accommodate rules, see for example, J. Rawls, "Two Concepts of Rules" in P. Foot (ed), *Theories of Ethics* (1967) p.144; R.G. Frey, "Act-Utilitarianism" in R.G. Frey (ed), *Utility and Rights* (1984), p.61. There are doubters: J.J.C. Smart, "Extreme and Restricted Utilitarianism" in Foot, *ibid*, at p.171.

⁵² *W v Egdell*, n.23 above, at 424.

⁵³ Cited in judgment of Sir Stephen Brown P, [1991] 2 F.L.R. 478, at 481. The Court of Appeal upheld the decision. See also *Gunn-Russo v Nugent Care Society* [2001] EWHC Admin 566; [2002] 1 F.L.R. 1.

⁵⁴ (1998) 25 E.H.R.R. 371 at para.104.

⁵⁵ *A Health Authority v X*, n.33 above, at 692; *Woolgar v Chief Constable of the Sussex Police* [1999] 3 All E.R. 604, at 606, 615; *R. (on the application of S) v Plymouth City Council* [2002] EWCA Civ 388; [2002] 1 W.L.R. 2583, at 2599; *Re R (a child)* [2004] EWHC 2085. Although an emphasis on restrictive disclosure forms part of the less political aspect of the test of necessity (see subsequent text) and therefore will be emphasised by the courts where they must test the necessity

Thus, despite the weight given to the freedom of the press and freedom of expression, there are strong utilitarian reasons why disclosures to the press should be treated more restrictively than disclosures for less publicly obvious purposes, even though the public interest justifying the latter might not be as strong as the public interest in free speech. The former are more likely to undermine the public interest protected by the obligation of medical confidence than the latter. It follows that the media cases are an unreliable guide to the capacity of the law to protect informational privacy in non-media cases.

The utilitarian rationale is not concerned with advancing informational autonomy. This might, at first, be considered to be incompatible with an approach informed by Article 8 jurisprudence, assuming that this requires domestic courts to accord more weight to informational autonomy.⁵⁶ However the European case law on medical privacy does not appear to equate informational privacy with informational autonomy but with utilitarian concerns.

Thus in *Z v Finland*, the European Court, while articulating the fundamental importance of protecting medical privacy, explained that:

“It is crucial not only to respect the privacy of a patient, but also to preserve his or her confidence in the medical profession and in health services in general. Without such protection those in need of medical assistance may be deterred from revealing such information of a personal and intimate nature as may be necessary in order to receive appropriate treatment and, even, from seeking such assistance thereby endangering their own health and, in the case of transmissible diseases, that of the community.”⁵⁷

The Court also commented that a failure to protect the privacy of medical information would affect not only the patient but would also undermine the community’s efforts to contain the AIDS pandemic.⁵⁸

These are consequentialist justifications which little to do with promoting patient autonomy. It is true that in relation to other aspects of Article 8, the Court has been more willing to identify the interests protected as being autonomy based or in the nature of dignitary interests.⁵⁹ Given the wide spectrum of situations falling under Article 8 and its open-ended nature, it may not follow that the principles articulated in dissimilar cases will be

of an interference but do not wish to engage on a more substantive review of it, this cannot wholly explain this legal development as it predates the Human Rights Act 1998: *W v Egdell*, n.23 above, at 416; *Re C (A Minor) (Evidence: Confidential Information)*, n.53 above, at 483.

⁵⁶ Phillipson and Fenwick, n.13 above, at pp. 662-663. G. Phillipson, “Transforming Breach of Confidence? Towards a Common Law Right of Privacy under the Human Rights Act” (2003) 66 MLR 726 at p.732. See also Case, n.5 above, at p.221.

⁵⁷ n.54 above, at para.95.

⁵⁸ *ibid*, at paras. 96 – 97. See also *MS v Sweden* (1999) 28 E.H.R.R. 313, at para.41.

⁵⁹ *Pretty v United Kingdom* (2002) 35 E.H.R.R. 1, at paras. 61 and 65; *Goodwin v United Kingdom* (2002) 35 E.H.R.R. 18, at para.90; *Glass v UK* (2004) 39 E.H.R.R. 14, at para.70.

recognised and applied in the context of medical privacy.⁶⁰ Nevertheless in *Gaskin v United Kingdom* and *MG v United Kingdom*, the Court recognised that denying the applicants access to information concerning them held by the authorities infringed their vital interest in receiving information necessary to know and to understand their identity.⁶¹ In *Peck v United Kingdom*, the release and use, without consent, of CCTV footage of the applicant attempting suicide in a public place, was held to be a breach of his Article 8 rights, including the right to identity and personal development, and the right to establish and develop relationships with other human beings and the outside world.⁶² These cases are significant because, although they made no reference to informational autonomy, they concerned, to some degree, the extent of individuals' ability to exercise control over information relating to them and therefore, arguably, aspects of the same, or closely connected, privacy rights protected by Article 8. If some aspects of this right to control were recognised as being based upon dignitary interests, a commitment to coherence in the law would justify recognising that the parallel right, in the context of medical confidentiality, is similarly underpinned. Furthermore in the recent decision of *Hannover v Germany*, a case concerning informational privacy and press intrusion, the European Court recognised that Article 8 protects interests in physical and psychological integrity and the development of personality.⁶³

(b) Mental integrity

As it happens the domestic courts have not differentiated between Article 8 cases which deal with informational privacy and those which concern other aspects of Article 8 when identifying the interests at stake in breach of confidence cases. The common law has therefore been more receptive to recognising dignitary interests, such as the protection of mental integrity, than the Strasbourg authorities on medical privacy require.

Thus in *X (a woman formerly known as Mary Bell) v SO*, an injunction was granted contra mundum to prevent the publication of the identity and whereabouts of Mary Bell and her daughter in order to protect Mary Bell's private interest in confidentiality and her Article 8 rights.⁶⁴ Butler-Sloss P interpreted Article 8 as covering physical and psychological integrity, a right to personal development and a right to establish and develop relationships with other human beings and the outside world.⁶⁵

⁶⁰ See D.J. Harris, M. O'Boyle and C. Warbrick, *Law of the European Convention on Human Rights* (1995), at p.303, who warn that the outcome of any particular case under Article 8 may not tell us much beyond its own facts.

⁶¹ *Gaskin v United Kingdom*, Series A (1989) 12 E.H.R.R. 29, at para.49; *MG v United Kingdom* (2003) 36 E.H.R.R. 3, at paras.27 – 29. See also *Rotaru v Romania* 8 BHRC 449, at para.43.

⁶² (2003) 36 E.H.R.R. 41, at para.57.

⁶³ [2004] E.M.L.R. 21, at [50] and [69].

⁶⁴ [2003] EWHC 1101; [2003] E.M.L.R. 37, at [38].

⁶⁵ *ibid.*, at [20], relying on *Botta v Italy* (1998) 26 E.H.R.R. 241, *Bensaid v UK* (2001) 33 E.H.R.R. 10 and *X v Netherlands* (1985) 8 E.H.R.R. 235, none of which concerned informational privacy. These interests have also been recognised in domestic cases concerning information access and disclosure not founded on confidence: *Re S (a child) (identification: restrictions on publication)* [2003] EWCA Civ 963; [2003] 3 W.L.R. 1425, at 1449-1450 and *R. (on the application*

These comments offer an instrumentalist interpretation of the values underlying Article 8. They do not necessarily offer the same basis of protection as a right to confidence based on informational autonomy. For example, where a patient opposes disclosure but it poses no risk to her mental integrity, there would be little reason to resist disclosure despite violating the patient's autonomy. Furthermore it is unclear whether a finding that these interests have been infringed is contingent on the individual suffering, or being at risk of, psychological harm as a result of disclosure, or whether mere distress would be sufficient. Butler-Sloss P., for example, stressed that the claimant would suffer substantial injury in the form of a recognisable psychiatric injury and her mental health problems would be exacerbated if the information was revealed.⁶⁶ Similarly in *Campbell v MGN*, Baroness Hale, who also acknowledged this rationale for protecting privacy, focused not only on Ms Campbell's distress at disclosure but also at the degree of physical and mental harm she would suffer as a result of her treatment being disrupted by the disclosures.⁶⁷

However, as the law has recognised that damages can be awarded in confidence for distress alone, the existence of a risk of psychological harm is likely to be more relevant when considering whether the right to confidence should be permitted to prevail over a public interest in disclosure, an issue considered in more detail subsequently.⁶⁸

(c) *Autonomy*

Until recently, the case law on confidentiality was remarkably silent regarding the principle of respect for autonomy in contrast to decisions on consent to treatment where the principle of respect for autonomy is often cited as the moral underpinning of the law.⁶⁹

Autonomy and confidentiality were first linked in *Douglas v Hello! Ltd (No. 1)*.⁷⁰ The facts are well known. Briefly, *Hello!* published unauthorised photographs of the celebrity wedding. The exclusive rights to photographs had been granted to *O.K.*, and the exercise was designed to spoil *O.K.*'s scoop. An action was brought against *Hello!* for breach of confidence. In the course of the hearing in the Court of Appeal for an interim injunction, Sedley L.J. indicated that the action for breach of confidence could be used to protect the claimants' privacy interests. He went on to comment that the law "can recognise privacy itself as a legal principle drawn from the fundamental value of personal autonomy."⁷¹ In his view the law of confidentiality

of Rose) v *Secretary of State for Health* [2002] EWHC 1593; [2002] 2 F.L.R. 962, at 972-973; *Re Angela Roddy (A Minor)* [2003] EWHC 2927; [2004] E.M.L.R. 8.

⁶⁶ [2003] E.M.L.R. 37, at [60].

⁶⁷ n. 36 above, at para.157. See also *Re S (A Child) (Identification: Restrictions on Publication)* [2003] EWCA Civ 963; [2003] 3 W.L.R. 1425, at 1450.

⁶⁸ For discussion of the case-law on damages for distress see references, n.34 above.

⁶⁹ *Re T (adult: refusal of medical treatment)* [1992] 4 All E.R. 649; *Re C (adult: refusal of medical treatment)* [1994] 1 All E.R. 819; *St Georges Healthcare NHS Trust v S* [1998] 3 All E.R. 673; *Re B v An NHS Hospital Trust* [2002] 2 All E.R. 449.

⁷⁰ [2001] Q.B. 967.

⁷¹ *ibid.*, at [126]. The statement has attracted much comment but this has been directed at the apparent recognition of privacy as a legal principle rather than at

protected privacy interests which in turn were derived from the principle of autonomy.

In the same year, *R. v Department of Health, ex p. Source Informatics Ltd* undermined claims that a connection could be made between the purpose of the law of confidence and the principle of respect for autonomy when the Court of Appeal ruled that there was no breach of confidence in the use of anonymised patient data.⁷²

Simon Brown L.J., while neither expressly accepting nor rejecting the proposition that in actions for breach of confidence a patient's autonomy interests were at stake, found that the law of confidence, as it applied to personal confidences, was concerned only to protect the confider's privacy. The disclosure of anonymised information did not invade privacy and so there was no breach of confidence. He also commented that, providing that a patient's privacy was safeguarded, a patient's will would not be thwarted if information obtained from the patient was used without consent, although this clearly does not follow.⁷³ The decision in *Source Informatics* illustrates that informational autonomy cannot necessarily be equated with informational privacy: the former may be violated even when the latter is not.⁷⁴

More recently in *Campbell v MGN*, Naomi Campbell successfully sued the *Mirror* for breach of confidence in relation to the publication of a story and photographs relating to her treatment for drug addiction.⁷⁵ Lord Hoffmann, together with Baroness Hale, acknowledged the principle of informational autonomy and emphasised that respect for autonomy required that people should have the right to control the dissemination of information about them, commenting that "the extent to which information about one's state of health, including drug dependency, should be communicated to other people was plainly something which an individual was entitled to decide for herself."⁷⁶

The significance of this judicial recognition of informational autonomy was reduced by a number of factors. Firstly the case involved disclosure of medical information by the press and, as discussed above, medical information receives a high degree of protection where such disclosures are concerned. It involves no radical change of approach to recognise the value of autonomy in this context. Secondly, Lord Hoffmann declined to explore

the link made between autonomy on the one hand and privacy and confidentiality on the other: see, for example Phillipson and Fenwick, n.13 above. Interestingly he has recently retracted the comment: Sir Stephen Sedley, "The Rocks or the Open Sea: Where is the Human Rights Act Heading?" (2005) 32 *Journal of Law and Society* 3, at p.16.

⁷² n. 33 above.

⁷³ *ibid*, at 440.

⁷⁴ Laurie, n.2 above, at 223-226. See also D. Beyleveld and E. Histed, "Betrayal of Confidence in the Court of Appeal" (2000) 4 *Medical Law International* 277.

⁷⁵ n. 36 above, Lord Nicholls and Lord Hoffmann dissenting.

⁷⁶ *ibid*, at para.53. See also *R. (on the application of Marper) v Chief Constable of South Yorkshire*, n.41 above, at 2217 (Baroness Hale); *Re Angela Roddy (A Minor)*, n.65 above; *Douglas v Hello (No. 3)* [2005] EWCA Civ 595, at paras.79 and 81.

the implications of recognising autonomy more generally and so the impact this case may have on other types of information disclosure, including disclosure by the NHS, is unclear. Thirdly, the other Law Lords did not adopt Lord Hoffmann's analysis. In contrast Lord Hope, in the majority, emphasised that the case raised no new principles.⁷⁷

Finally, the test for breach of confidence adopted by the other judges did not reflect the value of autonomy. The Law Lords endorsed an objective test in relation to both whether information could be considered private and whether there had been a breach of privacy through the use of such information. Information will be considered private if the person receiving it knows or ought to know it is fairly and reasonably to be regarded as confidential.⁷⁸ An invasion of privacy will be established if this information was disclosed when the person to whom it related had a reasonable expectation of privacy.⁷⁹ At times the tests were combined.⁸⁰ It might be argued that these elements amount to the same thing and that there is no need to distinguish them. Lord Hope commented that if information is obviously private the person to whom it relates would have a reasonable expectation of privacy.⁸¹ However this may not follow. Medical information is obviously private but a court may not always take the view that a person can reasonably expect it to remain undisclosed. For example, in *Source Informatics*, Simon Brown L.J. did not consider that individual objection to the use of private information for a broad range of NHS purposes would be reasonable and inclined to the view that, in relation to such uses, no breach of confidence could arise.⁸²

Informational autonomy would have been better respected if the test of whether there had been an invasion of privacy through the disclosure of private information had been subjective, based on an individual's actual expectations, even if the meaning of private information was objectively defined.⁸³ The unreasonableness of expectations regarding information use would then have been relevant in considering whether an invasion of privacy was justified.

In conclusion, informational autonomy is inadequately protected by the law. While this may be due to the limited judicial recognition it has received, a

⁷⁷ *ibid*, at 480.

⁷⁸ *ibid*, at 465 (Lord Nicholls).

⁷⁹ *ibid*, at 466 (Lord Nicholls).

⁸⁰ *ibid*, at 480 (Lord Hope), at 495 (Baroness Hale).

⁸¹ *ibid*, at 483.

⁸² n. 33 above, at 443-444. Although he also considered that the test of whether a use was reasonable should be judged by reference to the conscience of a reasonable person in the shoes of the recipient of the information, an approach that was rejected by the House of Lords. Nevertheless, the adoption of the objective test leaves patient privacy vulnerable to a court's views of what a reasonable patient would regard as unobjectionable use instead of being governed by the views of the individual affected.

⁸³ A similar point is made by Phillipson and Fenwick, n.13 above, at p.674. Moreham makes the additional point that the test focuses on whether the claimant can *expect* to have their privacy respected rather than whether they think they *should* have their privacy respected. He asks whether the parents of a missing child have a reasonable expectation that the media will not doorstep them: N. Moreham, "Recognising Privacy in England and New Zealand" (2004) *C.L.J.* 555.

more serious obstacle to its protection lies in the manner in which the domestic and European courts assess whether disclosure is justified in the public interest. The next sections will examine how the courts have approached this task and the problems they have met in doing so.

THE PUBLIC INTEREST DEFENCE AND ARTICLE 8(2)

The Scope of the Defences to Disclosure

In determining whether a disclosure of confidential information is justifiable the courts must take into account both the common law defence of the public interest and assess whether the justification falls within the grounds set out in Article 8(2). This permits disclosures which are in accordance with the law and are “necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals or for the protection of the rights and freedoms of others.”

It is unclear how the introduction of Article 8(2) will impact on the development of the common law of confidence. One possible consequence is that the public interest defence will develop to mirror the permissible grounds for disclosure in that provision, which are probably broader than the previously recognised common law public interest grounds.⁸⁴ For example Article 8(2) does not confine disclosures for the prevention of crime to serious crime and the ground of disclosure for the protection of the economic well being of the country is novel. Again, the use of information for research purposes could conceivably be justified by reference to the Article 8(2) ground of the protection of health even though it is unclear that the public interest defence previously encompassed such use.⁸⁵ This correlation between the common law and Article 8 (2) may not necessarily occur: Feldman for example considers that a divergence between the public interest defence and Article 8(2) is possible since the courts are not bound by Strasbourg jurisprudence but need only take account of it.⁸⁶ However what is significant is that the incorporation of Article 8(2) does not require the courts to give greater protection to informational privacy than previously, with one exception: insofar as an open-ended public interest defence of just cause or excuse had been recognised, the courts must now avoid interpreting it to be broader than Article 8(2), as interferences falling outside the terms of Article 8(2) would breach Article 8.⁸⁷ Given the breadth of Article 8(2), this should cause no great difficulties.⁸⁸

⁸⁴ Brazier, n.20 above, at p.81.

⁸⁵ Case, n.5 above, at pp.219-221.

⁸⁶ D. Feldman, “Information and Privacy” in J. Beatson and Y. Cripps (eds.), *Freedom of Expression and Freedom of Information* (2000), 299 at p.318. See also *Douglas v Hello (No. 3)* [2005] EWCA Civ 595 at para.53: the courts will “take account of Strasbourg jurisprudence”.

⁸⁷ R. Clayton and H. Tomlinson, *The Law of Human Rights* (2000), at paras.12.168-12.170. Lord Scott makes this point in relation to Article 10(2): *R. (on the application of ProLife Alliance) v BBC* [2004] 1 A.C. 185, at 242.

⁸⁸ Fenwick and Phillipson argue that in cases of media disclosure the courts have applied an over-broad test of the public interest which goes beyond the disclosure of an iniquity to justify disclosure of facts which are merely interesting to the

THE BALANCING EXERCISE: MORE STRUCTURE, MORE PROTECTION?

Nevertheless Feldman has argued that the need to justify interferences by reference to Article 8(2) may improve information protection because Article 8(2) is more scientific and certain than the common law and will require public authorities to reflect more carefully on the way rights and interests are balanced against each other.⁸⁹ In contrast to the under-articulated nature of the public interest defence, Article 8(2) requires a systematic approach to the resolution of conflicts between the public interest in disclosure and the protection of medical confidentiality. Thus the European Court applies a series of tests asking was the interference: (i) in accordance with the law; (ii) in pursuit of an aim listed in Article 8(2); and (iii) necessary in a democratic society. In relation to the last requirement it must be shown that the interference: (a) met a pressing social need (b) was supported by relevant and sufficient reasons and (c) was proportionate to the aim pursued.⁹⁰ In order to establish that an interference is proportionate, it must be shown that the aims pursued could not have been achieved in a less invasive manner, that the rights of the individual were taken into account and that adequate safeguards against abuse are in place.⁹¹

The case law following the HRA, however, illustrates a dilution in the strength of judicially accepted justifications for disclosure. In *A Health Authority v X* the Court of Appeal upheld an order for disclosure of medical records to a health authority, despite patient opposition, to enable the authority to investigate whether to initiate disciplinary proceedings against a general practitioner.⁹² The court found that there was a public interest in the proper administration of professional disciplinary hearings analogous to the established public interest in the due administration of criminal proceedings.⁹³ As this was an extension of the scope of this public interest, the protection afforded to confidentiality was weakened.⁹⁴ In reaching its conclusion the court relied upon *Woolgar v Chief Constable of the Sussex Police*. However in this case disclosure to the regulatory authorities was necessary for public protection reasons and there was evidence of a real risk to the public: a patient in a nurse's care had died and following a police

public. However, as they note, in media cases no clear distinction has been drawn between the public interest in press freedom and Article 10 rights, and other grounds for disclosure in the public interest. It would be unsafe therefore to cite these as authority for a broad public interest defence in non-media cases: "National Irish Bank v RTE and Finding the Balance: Breach of Confidence Privacy and the Public Interest Test in England and Ireland" (2004) 55 *NILQ* 118.

⁸⁹ n. 86 above, at pp.317, 324.

⁹⁰ *W v United Kingdom* Series A 121 (1988) 10 E.H.R.R. 29, at para.60; *Sunday Times v United Kingdom (No 1)* (1979-80) 2 E.H.R.R. 245, at para.62.

⁹¹ *Campbell v United Kingdom* Series A 233-A (1993) 15 E.H.R.R. 137, at para.48; *Klass v Germany* Series A 28 (1979-80) 2 E.H.R.R. 214, at para.55.

⁹² n. 47 above. Patient opposition was noted at first instance: *A Health Authority v X*, n.33 above, at 677.

⁹³ n.47 above, at 786.

⁹⁴ A similar approach was adopted in *Re B* [2004] EWHC 411; [2004] 2 F.L.R. 142, at 183-184. Para.2 of the Schedule to the Data Protection (Processing of Sensitive Personal Data) Order 2000, (SI 2000/417), permits disclosure for regulatory purposes if in the substantial public interest.

investigation, the police considered it necessary to pass on the information.⁹⁵ Again, in the leading case of *W v Egdell*, disclosure of a psychiatric report in breach of medical confidentiality was justified to avert a real risk of death or injury to the public if the information was not disclosed.⁹⁶ More recently there have been suggestions that disclosure can proceed where there is little evidence of risk. In *Re A (Disclosure of Medical Records to the GMC)*, a case involving the production of a child's medical records, Cazalet J. commented:

“If . . . the documents sought are only medical records . . . and the court is satisfied that these are or may be relevant to the GMC carrying out its statutory duties to protect the public against possible medical misconduct, it is hard to see what grounds, if any, the parents or other parties concerned may successfully raise against any such disclosure.”⁹⁷

These comments suggest that disclosure will be permitted, regardless of the nature or degree of risk to the public. Although the case predated the Human Rights Act 1998, in *A Health Authority v X* at first instance, Munby J. endorsed and applied these comments.⁹⁸ He considered that disclosure was permissible on the Article 8 (2) grounds of protecting public safety, health or morals, and the rights and freedoms of others.⁹⁹ However he made clear that he did not know whether there was any substance to the allegations under investigation.¹⁰⁰ There was no discussion of the degree of risk involved, nor evidence that there was a real risk. It was sufficient that the allegations were of a serious nature and that there were adequate safeguards in place against unauthorised disclosure.¹⁰¹

Munby J.'s approach is consistent with Strasbourg jurisprudence. For example, although *Z v Finland* is often cited as demonstrating the high priority accorded to the protection of medical information, *Z*'s private interest in privacy only prevailed over the public interest in relation to one of her complaints. Her husband, who was HIV positive, had been charged with the attempted manslaughter of women with whom he had had sexual relations. Disclosure of *Z*'s medical records to the prosecution, and witness orders made against her doctors, were found to be justified as a proportionate response to the pursuit of legitimate aims, namely the protection of the rights

⁹⁵ n. 55 above, at 615.

⁹⁶ n. 23 above, at 416, 424-425. Applied in *R v Crozier* [1991] Crim. L.R. 138 (a real physical risk to the safety of the public if a medical report was not disclosed). See also *Re L (Care Proceedings: Disclosure to Third Party)* [2000] 1 F.L.R. 913 (evidence indicated that the healthcare professional presented a real risk to patients).

⁹⁷ [1998] 2 F.L.R. 641, at 644.

⁹⁸ n. 33 above, at 691.

⁹⁹ *ibid.*, at 690.

¹⁰⁰ *ibid.*, at 676.

¹⁰¹ *ibid.*, at 697. The allegations concerned serious dispensing of medicine, inappropriate delegation of clinical responsibility and the performance of a sterilisation without proper consent: 690. See also *A v General Medical Council* [2004] All E.R. (D) 246 at paras. 136-137 but contrast the approach to risk in *R v A Local Health Authority, ex p. LM* [2000] 1 F.C.R. 736.

and freedoms of others and the prevention of crime.¹⁰² She succeeded in establishing that the disclosure of her name and medical condition in the domestic judgment was an unjustified interference with her right to medical privacy but this was because it did not serve any public interest at all, rather than because her private right outweighed a public interest in disclosure.¹⁰³ Disclosure of material from the court proceedings identifying Z and her medical condition ten years thence was a disproportionate response to the legitimate goal of ensuring that the administration of justice was transparent. However, even then, the possibility that the information could be legitimately disclosed at a later date was left open.¹⁰⁴

In *TV v Finland* the disclosure of the HIV positive status of a prisoner to prison staff directly concerned with his care was considered justified on the basis of protecting the rights and freedoms of others. The test of necessity was satisfied by demonstrating that the staff were bound by a strict obligation of confidentiality.¹⁰⁵

In *MS v Sweden*, Article 8(2) was again successfully advanced to justify the disclosure of sensitive medical information. MS had made a claim for industrial injury compensation from the Social Insurance Office (“SIO”). Without her knowledge her medical records, revealing that she had had a termination some time after her work related injury, were forwarded to the SIO. This interference with her rights was found to be justified because the information obtained was potentially decisive in the granting of public funds and therefore its communication was aimed at protecting the economic well being of the country. Weakly substantiated public interest arguments prevailed even though the harm done to MS by disclosure was arguably graver than the potential harm to the community.¹⁰⁶ Again, when determining the necessity of the interference, the Court focussed on the existence of adequate safeguards against disclosure.

These cases suggest that Article 8(2) does not present a particularly difficult obstacle for public bodies wishing to access confidential medical records. Partly this is due to the breadth of Article 8(2), which means that it is not hard for a public body to link an interference to one of the grounds contained therein, but partly it is linked to the restrictive manner in which the test of the necessity of interference has been applied. As Dickson has pointed out, in applying the test, the European Court has frequently focussed on procedural rather than substantive issues.¹⁰⁷ Of those cases which relate to informational privacy, the majority succeed because the interferences are not in accordance with the law or there are inadequate safeguards in place against abuse rather than, for example, that the social need to which the interference was directed is insufficiently pressing.¹⁰⁸ If these elements of the test are satisfied, it is

¹⁰² n. 54 above, at para.110.

¹⁰³ *ibid.*, at para.113.

¹⁰⁴ *ibid.*, at para.112.

¹⁰⁵ 76A D.R. 140 (1994).

¹⁰⁶ (1999) 28 E.H.R.R. 313, at paras. 42-44.

¹⁰⁷ B. Dickson, *Human Rights and the European Convention* (1997), p.78.

¹⁰⁸ These cases fall under respect for correspondence as well as respect for private life. See *Malone v UK* Series A No 82, (1984) 7 E.H.R.R. 14; *Kruslin v France* Series A No 176-A, (1990) 12 E.H.R.R. 547; *Herczegfalvy v Austria* Series A No

more difficult to establish an unjustified breach of Article 8¹⁰⁹ unless the interference has had a substantial impact on the individual, interfered with the essence of the relevant right, or interfered with other protected Convention rights and important public interests.¹¹⁰

The use of information by the NHS is unlikely to fall foul of this restrictive application of the test of necessity. Pattenden has suggested that a broad public interest defence to the action for breach of confidence may breach the requirement that interferences must be in accordance with the law.¹¹¹ A law must be sufficiently detailed to enable individuals to regulate their conduct to avoid breaching it and the public interest defence may not satisfy this.¹¹² On the other hand the European Court has recognised that laws may be couched in vague terms to avoid rigidity and the limits of the defence are likely to be set by Article 8(2) itself.¹¹³ Furthermore, the safeguards in place against abuse are almost certainly adequate. Disclosures will be regulated by the Data Protection Act 1998, by the law of confidence, by professional obligations of confidence and possibly by the criminal law.¹¹⁴ Again, it is unlikely that the use of information for NHS purposes, such as auditing or management, would have a significant impact on most patients and, apart from the public interest in medical confidentiality, no other weighty public interest or Convention right would normally be affected by disclosure.¹¹⁵ The cases demonstrate that medical privacy does not receive a high degree of protection when demands for disclosure are made by public bodies citing the public interest and/or one of the Article 8(2) grounds for disclosure. Article 8(2) may have required a more structured approach to considering whether disclosures of private information can proceed in the public interest but it has

242-B, (1992) 15 E.H.R.R. 437; *Halford v UK* (1997) 24 E.H.R.R. 523; *Huwig v France* Series A No 176-B, (1999) 12 E.H.R.R. 528; *Perry v UK* (2004) 39 E.H.R.R. 76. See also *R. (on the application of Szuluk) v Governor HMP Full Sutton* [2004] EWHC 514; [2004] A.C.D. 45: inappropriate application of blanket rule intercepting medical communications. The scope of the right to confidentiality is broader than that of respect for correspondence however: *AD v Netherlands* 76A D.R. 157 (1994).

¹⁰⁹ *Leander v Sweden* (1987) 9 E.H.R.R. 433; *Chare (nee Jullien) v France*, n.41 above; *Klass v Germany*, n.91 above; *TV v Finland*, n.105 above; *MS v Sweden* n.58 above.

¹¹⁰ For example, intercepting detainees' correspondence strikes at the essence of the right to respect for correspondence, given their reliance on such correspondence to communicate with the outside world: see *Herczegfalvy v Austria*, n.108 above, at para.91. Interception of lawyer/client correspondence is also closely scrutinised due to the impact on the public interest in the due administration of justice, legal professional privilege and Article 6 rights: *Niemietz v Germany* Series A No. 251-B, (1992) 16 E.H.R.R. 97.

¹¹¹ n.7 above, at para.11.07. Feldman thinks it unlikely: n.86 above, at p.317.

¹¹² *Doerga v Netherlands* (applic.50210/99, 27 April 2004).

¹¹³ *Silver v UK* (1983) 5 E.H.R.R. at para.88.

¹¹⁴ The Health and Social Care (Community Health and Standards) Act 2003, s.136. Such safeguards were held to be adequate in *MS v Sweden*, n.58 above and *Chare (nee Jullien) v France*, n.41 above.

¹¹⁵ *Venables v News Group Newspapers Ltd* [2001] 1 All E.R. 908 and *Carr v Newsgroup Newspapers Ltd* (unrep) Q.B.D 24 February 2005 illustrate an exception to this as Article 2 rights were at risk.

not required a more stringent one. As a result informational autonomy is not well protected, even incidentally.

Assessing the Public Interest in the Disclosure of Medical Information: Constitutional Problems

McHarg has argued that the failure of the courts to give adequate protection to Convention rights is a result of the difficulties the judiciary have had in finding a politically defensible method of reconciling rights and the public interest. She argues that one model of adjudication adopted by the European Court involves balancing rights, conceived of as protected interests, against the public interest, conceived of as a common interest (being the interests which people in a society have in common), with the stronger prevailing. The Convention provides no objective criterion against which these different kinds of interest can be weighed and judges may not legitimately construct their own, as they are not in a position to assess the importance of a public interest, nor what serves the public interest. Only the public can do so through political institutions. The courts are limited to enquiring whether an interference is actually necessary to achieve a particular public interest. If it is, judges cannot prohibit it. To do so would result in the undemocratic and illegitimate substitution of judicial views of the appropriate balance between rights and collective goals for that of the State and its emanations.¹¹⁶ McHarg's argument highlights the problems the judiciary may confront when a right conflicts with a public interest and explains why, in such cases, the courts may choose to focus on the more procedural and less political aspects of the Article 8(2) test.

McHarg suggests that her analysis is relevant to courts at both European and domestic levels although the problem may be less acute domestically owing to the non-applicability of the margin of appreciation.¹¹⁷ This permits states some leeway in striking the balance between rights and public interests within their territories as a reflection of respect for national sovereignty and because national authorities are better placed than the Court to assess what is necessary in the national context. It has no place in domestic jurisprudence.¹¹⁸ Nevertheless the domestic courts do permit public authorities a discretionary area of judgment in making decisions about whether the public interest should be advanced at the expense of rights and this is partly due to a concern about the democratic legitimacy of judicial

¹¹⁶ A. McHarg, "Reconciling Human Rights and the Public Interest: Conceptual Problems and Doctrinal Uncertainty in the Jurisprudence of the European Court of Human Rights" [1999] 62 *MLR* 671, at pp. 676-680. McHarg identifies a second model of adjudication which views rights as rules with the public interest goals forming exceptions to those rules. A right will be upheld if an interference affects its core purpose. The problem of democratic legitimacy arises in this model also in relation to the Court determining the purpose of the right and in addressing interferences which affect it only peripherally: at pp.680-683. Both models are evident in the Article 8 jurisprudence under discussion but the balancing model dominates in the context of medical information.

¹¹⁷ *ibid.*, at pp.695-696.

¹¹⁸ *R. v DPP, ex p. Kebilene* [2000] 2 A.C. 326 at 380 (Lord Hope).

intervention.¹¹⁹ Furthermore the courts exercise greater deference where, as with Article 8, the right is qualified.¹²⁰

The argument that it is undemocratic or illegitimate for the judiciary to permit rights to override the public interest is contested.¹²¹ Greer, for example, argues that the Convention, interpreted teleologically, produces a number of principles of interpretation including the rights principle which requires that in a democratic society Convention rights should be protected by national courts and the priority to rights principle, which requires that rights should be privileged over collective goods.¹²² This does not mean that rights should be treated as trumps, in the sense that they must nearly always prevail over collective welfare.¹²³ As Ashworth points out, this interpretation is excluded by the structure of Article 8 itself which permits intervention on a broad range of grounds provided that it is necessary in a democratic society.¹²⁴ However it does mean that it can be legitimate for the courts to permit a right to prevail over the collective interest.

It is suggested, therefore, that arguments about lack of democratic mandate need not prevent the courts from permitting informational privacy to prevail over public interests in disclosure. However McHarg's argument highlights problems relating to what Jowell has termed the "institutional capacity" of the courts.¹²⁵ The courts' ability to privilege rights over collective interests will be limited where the assessment of the public interest lies outside their expertise and within the specialist knowledge of the public bodies concerned or when the evidence the courts would have to review to assess the weight of

¹¹⁹ *Secretary of State for the Home Department v Rehman* [2001] UKHL 47; [2003] 1 A.C. 153, at 195 (Lord Hoffmann); *R. (on the application of Holding & Barnes Plc) v Secretary of State for the Environment, Transport and the Regions* [2001] UKHL 23; [2003] 2 A.C. 295, at 325-327 (Lord Hoffman); *International Transport Roth GmbH v Secretary of State for the Home Department* [2002] EWCA Civ 158; [2003] Q.B. 728, at 765 (Laws L.J.); *R. (on the application of ProLife Alliance) v BBC*, n.87 above, at 256 (Lord Walker). See also J. Waldron, "A Right-Based Critique of Constitutional Rights" (1993) 13 *OJLS* 18; J. Allan, "Bills of Rights and Judicial Power-A Liberal's Quandary" (1996) 16 *OJLS* 337; J.A.G. Griffith, "The Common Law and the Political Constitution" [2001] 117 *LQR* 42.

¹²⁰ *R. v DPP, ex p. Kebilene*, n.118 above, at 381 (Lord Hope); *International Transport Roth GmbH v Secretary of State for the Home Department*, n.119 above, at 766 (Laws L.J.); *R. (on the application of ProLife Alliance) v BBC*, n.87 above, at 256 (Lord Walker); *Brown v Stott* [2003] 1 A.C. 681, at 720 (Lord Hope).

¹²¹ A. Ashworth, *Human Rights, Serious Crime and Criminal Procedure* (2002), at pp.71-72; F. Klug, "Judicial Deference under the Human Rights Act 1998" (2003) *EHRLR* 125; J. Jowell, "Judicial Deference: Servility, Civility or Institutional Capacity?" [2003] *PL* 592; A. Lester, "The Human Rights Act 1998 – Five Years On" (2004) *EHRLR* 258. Lord Steyn, "Deference: a Tangled Story" (2005) *PL* 246. It should be noted that McHarg does not necessarily subscribe to this argument but rather identifies it as following from the balancing model of adjudication adopted by the European Court of Human Rights.

¹²² S. Greer, "Constitutionalizing Adjudication under the European Convention on Human Rights" (2003) 23 *OJLS* 405, at pp.410-414.

¹²³ As argued by R. Dworkin, *Taking Rights Seriously* (1977) Ch.4.

¹²⁴ n.121 above, at pp.75-77.

¹²⁵ n.121 above, at p.598.

a public interest raises issues of justiciability, such as evidence over resource allocation.¹²⁶ In such cases the courts will be less interventionist in scrutinising interferences with private rights.

In addition, McHarg points out that although designating an interest as a right is generally understood as signifying that it can potentially override collective aims, the balancing exercise adopted by the European Court robs rights of their peremptory status.¹²⁷ Rights have been treated as simply a factor in the balancing exercise which has resulted in them being “balanced away”.¹²⁸

INFORMATIONAL AUTONOMY, HEALTH INFORMATION AND THE PUBLIC INTEREST

In the light of the above it might be queried whether the incipient judicial acknowledgment of the interests of informational autonomy and mental integrity will result in better protection for medical privacy particularly in relation to uses of information by the NHS for, for example, administrative, audit and management purposes. These may be justified as protecting the economic well being of the country. To reject an assertion that such disclosures should take place in the public interest, the courts would have to consider evidence concerning cost and policy content which may be regarded as non-justiciable.¹²⁹ Even though there has been some indication that, as a result of the Human Rights Act 1998, the courts may be prepared to consider such evidence, the courts will not carry out a full merits review of executive and administrative decisions. As a result the extent to which that evidence will be addressed is limited.¹³⁰ Again, the courts will exercise greater deference where the aims pursued through an interference are socio-economic.¹³¹ Assessments of such public interests arguably fall outside the institutional capacity of the courts. Although there has been some indication that the courts may subject policy decisions to increased scrutiny, this has generally taken place in the course of judicial review, within the parameters of a legislative framework which has set limits on the extent to which judges

¹²⁶ *R. v Cambridge Health Authority, ex p. B* [1995] 2 All E.R. 129, at 136; *R. (on the application of Pfizer Ltd) v Secretary of State for Health* [2002] EWCA Civ 1566; [2003] 1 C.M.L.R. 19, at para.17.

¹²⁷ n.116 above, at p.680.

¹²⁸ See also Ashworth, n.121 above, at p.65 who perceives a “balancing away” of Article 6 rights by domestic courts.

¹²⁹ See cases n.126 above; *R. (on the application of ProLife Alliance) v BBC*, n.87 above, at 240 (Lord Hoffman). The extent to which the courts should treat such issues as non-justiciable is debated: D. O’Sullivan, “The Allocation of Scarce Resources and the Right to Life under the European Convention on Human Rights” [1998] *PL* 389; K. Syrett, “Impotence or Importance? Judicial Review in an Era of Explicit NHS Rationing” (2004) 67 *MLR* 289; Lord Steyn, n.121 above.

¹³⁰ See, for example, the consideration of statistics and policy arguments in support of the Road Traffic Act 1988, s.172 in *Brown v Stott*, n.120 above, at 704-705 (Lord Bingham), at 709-711 (Lord Steyn).

¹³¹ *R. v DPP, ex p. Kebilene*, n.118 above, at 381; *Wilson v First County Trust Ltd* [2004] 1 A.C. 816, at 844 (Lord Nicholls); *Evans v Amicus Healthcare Limited* [2004] EWCA Civ 727; [2004] 2 F.C.R. 530, at para.65. See R.G. Lee, “Resources, Rights, and Environmental Regulation” (2005) 32 *Journal of Law and Society* 111, at pp.123-125.

can intervene and provides external guidance on when intervention is justified.¹³² The action for breach of confidence contains no such parameters and, given this lack of guidance, the courts may exercise caution in intervening when the grounds for disclosure of patient data are socio-economic. If the right to confidentiality was prioritised over such public interest aims, this could have unexplored ramifications for the State and the NHS in terms of cost, allocation of resources, and patient safety and could have a disruptive impact on the running of an enterprise which pursues democratically legitimated social goals.¹³³

However the recognition of informational autonomy or mental integrity may make a difference in other types of case. For example, the disclosure of a competent elderly patient's medical information to social services, without consent, to enable social services to offer her assistance, might be justifiable on the Article 8(2) grounds of the protection of health. This is not a true collective interest but rather an articulation of the patient's welfare interests which the courts are well placed to assess. Where the interest underpinning the right to medical confidentiality is identified as the public interest, the case law examined earlier suggests that it is unlikely that a court would consider that disclosure impaired this interest. It would be relatively unproblematic therefore to find the disclosure justified and a court might wish to do so to avoid stigmatising a well intentioned healthcare professional. If the right is based on the protection of mental integrity and the patient did not suffer psychological harm as a result of disclosure, the same conclusion follows. In these circumstances, the principle of informational autonomy provides both the only and, arguably, an overriding reason for upholding confidentiality.¹³⁴

Between these two extremes lies the use of information for regulatory purposes. As seen, this has been justified on the basis of public safety and the due administration of justice. It lies well within the institutional capacity of the courts to assess these public interests and the impact of permitting a right to medical confidentiality to override the public interest. A close degree of scrutiny can therefore be exercised so, for example, the courts should be satisfied that a disclosure was the least invasive interference possible in pursuit of the public interest goals. Regulatory bodies should only be able to obtain medical records where there is a real risk of injustice

¹³² See, for example, *R. v North East Devon Health Authority, ex p. Coughlan*, [2001] Q.B. 213: resource allocation quashed after reviewing provisions of the National Assistance Act 1948 and the National Health Service Act 1977.

¹³³ As set out in the National Health Service Act 1977. In *A Health Authority v X*, although the Munby J. dismissed the argument that the courts should not be able to examine and veto the uses of health information by a health authority, he was, in fact, non-interventionist, permitting disclosure with little attention to the need for the disclosure: n.33 above, at 696-697.

¹³⁴ By analogy with the law on consent to treatment, respect for autonomy bars action in a competent patient's best interests without that patient's consent: n.69 above. The current legal position on disclosures in a competent patient's best interest is unclear. It is supported by some commentators, judges and by professional guidance: *Cornelius v De Taranto*, n.39 above, at para.62; *H (A Healthcare Worker) v Associated Newspapers Limited*, n.47 above, at para.28; J.K. Mason, R.A. McCall-Smith and G.T. Laurie, *Law and Medical Ethics* (2002), at p.244; GMC, n.32 above, at para.27.

or public harm if the conduct of the healthcare professional is not fully investigated.¹³⁵ The precise degree of risk required would vary depending upon the conception of the nature of the interest underpinning the right and the impact of the interference upon this interest. If informational autonomy was recognised, it would be acknowledged that all disclosures would interfere with the underlying interest and stronger evidence of risk would have to be advanced than if no interference was recognised. At the same time, if disclosure posed a risk to mental integrity, additional evidence of risk should be required given the potential impact of disclosure upon the individual.

CONCLUSION

Information technology developments in the NHS, including the introduction of electronic patient records, will make it easier for those who wish to access patient data to do so, whether for public interest purposes or for less beneficent reasons. The data will be vulnerable to access by a larger number of persons and it will be more complete and detailed than at present.¹³⁶ The practical barriers to information access, which incidentally assisted in protecting patient privacy, are being eliminated. At the same time there is little legal resistance to disclosure where disclosure is sought by public bodies in the public interest. The combination of these developments pose a substantial risk to patient privacy, at least where information is sought by public bodies. However, given the minimal impact that most disclosures would have on individual well-being and the community need for efficient public services, the Government argues that citizens have a responsibility to yield control over personal information in the public interest.¹³⁷ It can be argued that a wrongly placed emphasis on informational autonomy will be harmful to the NHS and to the public good. These arguments must be treated with caution

Those who argue for greater respect for patient autonomy would agree that rights are not absolute and must yield to the public good. The problem is how the balance should be struck and by whom. At present the law fails to give adequate weight to the privacy rights at risk. In particular, if rights are to be overridden, there should be credible evidence of harm if they are upheld. Further, it is not enough to override the right to medical privacy that there will be some cost to society through non-disclosure: the cost must be significant.¹³⁸ Despite this, disclosure has been obtained where little evidence of harm to the public interest has been forthcoming.

¹³⁵ A view supported by Brazier, n.20 above, at pp.63-68.

¹³⁶ An example of the risk posed to patient privacy by computerised records can be found in *Ashworth Hospital Authority v MGN Ltd* [2001] 1 W.L.R. 515 where evidence was given that over 200 people could access patient files held on the hospital's database: at 522. Notably this case involved the leak of patient data to a national newspaper.

¹³⁷ PIU, n.10 above, at para.5.04. For an analysis of the Government's approach see Perri 6, Raab and Bellamy, n.4 above, at p.116. It is shared by some of the judiciary: see Lord Steyn in *Brown v Stott*, n.120 above, at 707-708; and some commentators: Hagger, Woods and Barrow, n.20 above, at pp.107-109. For a critique of this approach see Ashworth, n.121 above, at pp.119-124.

¹³⁸ Dworkin, n.123 above, at p.200.

In order to safeguard privacy, patients must be granted greater control over their records. There are a number of ways this could be done, each with differing financial implications. The most expensive, and possibly most impracticable, would be to require express consent to all information uses.¹³⁹ A more limited exercise of informational autonomy would be to allow patients to opt out from information disclosure.¹⁴⁰ It has been argued that even this would be undesirable as it could corrupt the results obtained from essential research and audit, resulting in both financial and non financial costs to society.¹⁴¹ However there is a dearth of empirical evidence to support arguments that permitting opt-out would be harmful. It seems unlikely that this would usually cause harm given that the majority have indicated that they would permit their information to be used within the NHS at least.¹⁴² Furthermore, the very technological advances which threaten patient privacy are likely to reduce the practical difficulties and cost of recording and communicating patient opt-out. In the future it will be possible to keep a central record, accessible to all healthcare providers, of whether patients wish to opt out from information sharing. Furthermore it is anticipated that it will be possible to separate particularly sensitive information from the rest of the patient records in an “electronic envelope” and to keep audit trails of when, and by whom, patient information is accessed.¹⁴³ In most cases therefore, the balance between respect for patients’ rights and the public interest is likely to require that patient opt-out is observed.

The question remains regarding who should determine when a patient’s choice to opt out should be overridden in favour of economic and functional arguments in favour of disclosure. It has been argued here that the courts, applying the common law of confidence, may not have the capacity to protect informational autonomy and privacy against demands for disclosure made on such grounds. Informational autonomy could receive better legal protection if there was the political will to legislate to grant patients a right to opt out of information disclosure which could only be overridden in limited circumstances. Recently the NHS and the Minister for Health issued a

¹³⁹ For a review of the different approaches including costs implications see Cambridge Health Informatics, *Consent-based Use of Patient Information-Input Assessment* (January 2001). However recent information technology developments may make it possible to obtain, record and communicate express consent cheaply: see n.143 below and accompanying text.

¹⁴⁰ An approach supported by the GMC, n.32 above, at paras.13 and 21. Opt-out arguably promotes the public interest by minimising the loss of data which might occur because of, for example, problems in seeking consent. It strikes a balance tilted in favour of disclosure whilst still respecting patient autonomy. The argument that opt-out respects patient autonomy is only valid, however, insofar as patients are aware of the uses to which their information is put and are aware that they can opt out and increasing patient awareness has proved problematic.

¹⁴¹ Hagger, Woods and Barrow, n.20 above, at pp.106-107.

¹⁴² n.9 above.

¹⁴³ For a more detailed account of these, and other, security measures see NHS Connecting for Health, *A Guide to the National Programme for Information Technology* (2005) pp.20-21 and generally: http://www.connectingforhealth.nhs.uk/all_images_and_docs/NPfit%20brochure%20Apr%2005%20final.pdf. (Last visited 10 July 2005).

Guarantee on the use of electronic patient records which promises that in future patients will be able to veto access to their electronic records save where disclosure is compelled or it is necessary to avoid harm to a third party.¹⁴⁴ However, while the Guarantee gives informational autonomy prominence, its practical effect is unclear. At present it amounts to no more than a statement of intent regarding future developments: how it will apply in practice is uncertain.¹⁴⁵ It has no legal effect and it grants patients no remedies if their opt out is disregarded.¹⁴⁶ Its influence on professional practice is also uncertain since it conflicts with the guidance provided by professional bodies, the NHS's own Code of Practice on Confidentiality, the common law and the DPA, all of which permit varying degrees of disclosure in the public interest.¹⁴⁷ For these reasons alone it is highly unlikely that the Guarantee can be regarded as an indication of legislative intent. Furthermore, the recent initiatives which have eroded patient privacy, and the Government's communitarian ethos, suggest that the protection of medical privacy and informational autonomy, at a cost to the public interest, may not be prioritised by the legislature in the foreseeable future.¹⁴⁸ The Government's attitude to the protection of medical privacy is ambivalent at best. It has not hesitated to legislate to set aside privacy where it has considered it necessary and so it is likely to avoid creating an additional legal obstacle to information disclosure by giving patient opt-out unequivocal legal force. It remains to be seen therefore whether the Government's rhetoric will translate into a legislative commitment to patient autonomy when the State wishes to access patient data to serve public interest purposes.

¹⁴⁴ NHS, *The Care Record Guarantee. Our Guarantee for NHS Care Records in England* (2005), p.4. para.3. There is however a significant ambiguity in the Guarantee. Elsewhere it provides that by 2008 patients will be able to veto access to sensitive parts of their records save where it would put third parties at risk or when disclosure is *allowed* by law: at p.7. The latter phrasing allows for discretionary disclosure for public interest purposes whereas disclosure on grounds of legal compulsion only does not cover public interest disclosures, nor disclosure under the numerous statutory gateways which exist. If patient veto will be overridden where the law permits, as opposed to requires, disclosure there will be very little change to the current position.

¹⁴⁵ The electronic patient record programme will begin to roll out in 2006.

¹⁴⁶ Although it promises that the NHS will take disciplinary action: n.144 above, at p.5, para.12.

¹⁴⁷ Department of Health, *Confidentiality NHS Code of Practice* (November 2003), Annex B, pp. 34-35; GMC, n.32 above.

¹⁴⁸ Perri 6, Raab and Bellamy, n.4 above. There is little sign that the trend to increase the number of bodies who can access identifiable patient data is on the wane: the Government has recently consulted on permitting the NHS Counter Fraud and Security Management Service access to patient data: see Department of Health, *Access to Relevant Documents, Records and Data to Counter NHS Fraud: A Paper Consultation* (October 2004) and *Report on Consultation* (May 2005) <http://www.dh.gov.uk/Consultations/ResponsesToConsultations/ResponsesToConsultationsDocumentSummary/fs/en> (last visited 16 August 2005). For criticism see also the Patient Information Advisory Group, *Response* (January 2005): <http://www.advisorybodies.doh.gov.uk/piag/counterfraud-accesspiagresponse.pdf> (last visited 16 August 2005).

BEYOND RATIONALITY? JUDICIAL REVIEW AND PUBLIC CONCERN IN THE EU AND THE WTO

*Chris Hilson, Professor of Law, The University of Reading**

Introduction

Within the literature on risk regulation, there are, broadly speaking, two camps. On the one hand, there are the “rationalists”, who place considerable emphasis on apparently objective methods such as scientific risk assessment and cost-benefit analysis.¹ Rationalists regard public perceptions of risk as “irrational” and decry the perverse effects that these can have on regulatory priorities. In their view, society allocates too many of its resources to regulating risks which the public perceives as significant but which science suggests are relatively minor; and not enough to risks which are proven killers but which the public perceives as much less significant. On the other hand, there are the populists,² who question the rationalist assumption that scientific risk assessment is wholly objective, and who argue that public perceptions – far from being irrational – are simply exhibiting an alternative rationality. From the populist perspective, public perceptions of risk are democratically important and should be factored into the regulatory decision-making process from the outset.

This article is not just about the public perception of risk however. It is about public concern more generally. With many risk activities, the public may well be concerned because they perceive a risk to their own health or to the environment despite the science suggesting that there is no risk. Nevertheless, they may also have concerns about the activity on broader ethical, religious or political grounds. With genetically modified organisms (GMOs), for example, the public may – despite the lack of clear scientific evidence of a risk – perceive GMOs as a health risk. However, they may also have religious concerns about humans “playing God” with nature, or political concerns about multinational biotech companies monopolising the food sector.

* I am grateful to Veerle Heyvaert, Maria Lee and an anonymous referee for their comments. All errors remain my own.

¹ The leading exemplar being Sunstein, *Risk and Reason: Safety, Law, and the Environment* (Cambridge University Press, 2002). See also Breyer, *Breaking the Vicious Circle: Toward Effective Risk Regulation* (Harvard University Press, 1993).

² See e.g. Winickoff, Jasanoff, Busch, Grove-White and Wynne, “Adjudicating the GM Food Wars: Science, Risk, and Democracy in World Trade Law” (2005) 30 *Yale J. Int’l L.* 81; ESRC Global Environmental Change Programme, *The Politics of GM Food: Risk, Science and Public Trust*, Special Briefing No. 5 (University of Sussex, 1999); Marris, Wynne, Simmons & Weldon, *Public Attitudes to Agricultural Biotechnologies in Europe* (2001), available at <http://www.lancs.ac.uk/depts/ieppp/pabe/>; Gillette and Krier, “Risk, Courts, and Agencies” (1990) 138 *U. Pa. L. Rev.* 1027.

The aim of this article is to examine the extent to which, under EU and WTO law, decision-makers are legally entitled to take public concern into account. An answer to this question will necessarily involve an examination of how the issue of public concern has been dealt with by the European Court of Justice and the dispute resolution bodies of the WTO, in reviewing the decisions or actions of states or supranational authorities.³ From an analysis of these issues, we should be able to consider the extent to which these judicial bodies have been willing or able to move beyond the apparent rationality of science to take public concern into account.⁴

EU Case Law on Public Concern

Before examining the cases themselves, it is worth noting that, in a number of them, public perception of risk or broader public concern is raised indirectly via the notion of consumer confidence. In essence, while science may have revealed no evidence of risk associated with a particular product, the public may nevertheless perceive there to be a risk or object to it on broader grounds and thus reject the product. From the point of view of public decision-makers wanting to support producers, it becomes a matter of attempting to restore consumer confidence in the relevant product – confidence or trust that has been dented by a perception of risk (albeit one that may not be objectively justified by way of scientific evidence) or conflict with wider values. Consumer confidence in these cases therefore acts as a proxy for public concern.⁵

Turning now to the cases, the issue of public concern was first raised in two key cases involving the well-known “Beef Hormones” controversy.⁶ In *UK v*

³ Although the respective cases are not precisely on all fours (in that, as we shall see, the WTO ones concern free movement of goods and the EU ones concern review of Community legislative action rather than Member State regulations impeding free movement), they both involve “judicial review” of political action. On the WTO from a judicial review perspective, see Scott, “European Regulation of GMOs: Thinking About ‘Judicial Review’ in the WTO”, Jean Monnet Working Paper No. 04/04, available at <http://www.jeanmonnetprogram.org/papers/04/040401.html>

⁴ The assumption made in the article is thus that science and public concern are *separate*, with the latter typically being used to refer to scientifically *unjustified* risk perceptions (on the basis that there is no evidence of risk) and/or other, *e.g.* political/ethical-type concerns. However, *N.B.* that public concerns can of course be scientifically *justified* and may be incorporated *within* scientific assessment (*e.g.* public framing of the risk or identification of risk pathways may be more extensive than those initially adopted by scientific experts). *N.B.* also that some in the populist camp claim that even less evidentially specific, more speculative public anxieties about remaining uncertainties are still, in a sense, scientific.

⁵ Of course consumer confidence cannot always act as a proxy for public concern, as the public may also be concerned about risks that are not closely related to a consumable product.

⁶ Which will be returned to below in the section on the WTO. Both the EC and WTO aspects of the saga are discussed in: Scott, “On Kith and Kine: Trade and Environment in the EU and WTO” in Weiler (ed.), *The EU, WTO and NAFTA: Towards a Common Law of International Trade* (OUP, 2000), p.125; Joerges, “Law, Science and the Management of Risks to Health at the National, European and International Level – Stories on Baby Dummies, Mad Cows and Hormones in

Council,⁷ the applicant claimed, first, that Council Directive 85/649 – which banned the use of certain growth-promoting hormones in livestock farming – was adopted using the wrong legal base. The UK claimed that the Directive – which was adopted solely on the basis of Article 43 (now 37) of the Treaty, requiring only qualified majority voting – should have been based on Article 100 (now 94), which required unanimity and which would therefore have enabled the UK to reject the measure. The aim of the Directive, according to the recitals in its preamble, was to protect human health and consumer interests with a view to eliminating distortions of competition and bringing about an increase in consumption of the product in question.⁸ While the UK effectively conceded that the Directive had a legitimate agricultural policy aim, it claimed that the protection of public health was the predominant aim and that this required recourse to Article 100.⁹ However, the Court rejected this argument, stating that the agriculture Article 43 legitimately included health objectives.¹⁰ In relation to this first argument, the case is of little direct interest from the point of view of public concern – other than as a counter-example to most of the cases considered below, where applicants have tended to argue not that the relevant measure *was* aimed at public health but rather that it *was not*. In this respect, the UK’s first argument also sits rather uneasily with its second claim considered below which, as we shall see, suggests that public health *was not* really the main objective.

The UK’s second claim was more directly linked to the issue of public concern – at least in the Advocate General’s Opinion, if not in the ruling of the Court itself. The UK claimed that the reasons for Directive 85/649 were insufficiently stated. According to the UK, a scientific report, drawn up as a result of a duty in an earlier Directive,¹¹ pointed against the need for a ban on scientific, health grounds. In its statement of reasons, Directive 85/649 had failed to advance any reasons for not taking the findings of that report into account. On this point, Advocate General Lenz began by observing that, while the Commission had to take the report into account in drafting its proposal for a directive, the Council was under no duty to do so.¹² In a passage that is worth quoting at length, he then went on to state that:

“Moreover, it is clear from the arguments put forward in the proceedings that, in view of the reaction of the Economic and Social Committee and of the Parliament to the first Commission proposal of June 1984, the predominant concern of the Commission, and thereafter of the Council as well, was

Beef” (2001) 7 Col. J. of Eur. L. 1; and Majone, “What Price Safety? The Precautionary Principle and its Policy Implications” (2002) 40 J.C.M.S. 89.

⁷ Case 68/86 *UK v Council* [1988] E.C.R. 855. For comment on this case, raising a number of the issues discussed here, see Joerges (n.6 above) pp.9-10; and Majone (n.6 above) pp.96-97.

⁸ Judgment, para.20.

⁹ See para.16 of the A.G.’s Opinion.

¹⁰ Para.12.

¹¹ Art.8 of Dir. 81/602, which stated that “the Commission shall submit to the Council a report on the experience acquired and scientific developments, accompanied, if necessary, by proposals which take these developments into account.”

¹² Para.38.

not so much the safeguarding of the health of consumers (a problem which the scientific group considered above all) but to take into account the interests of consumers in general (*since it could be seen that meat from animals treated with hormones is widely rejected*). That attitude is without question one of the matters falling under “experiences acquired” referred to in Article 8 of Directive 81/602. Once the emphasis of the intended measure centred on that aspect, however, there was really no reason to examine the health problem in particular, which had evidently been the main focus of interest in the past, and so the fact that in the preamble to the contested directive the Council did not go into the partial findings of the scientific group, which the Commission did make available to the Council, certainly cannot be regarded as a failure to state reasons.”¹³

Here he is stating, in effect, that because the measure was mainly driven by consumer concern, it did not matter that the Council failed to refer to a scientific report stating that there was no risk to health. The suggestion is that, whatever scientists said as to the safety of hormones in meat in such a report, consumers would continue to perceive a health risk, or object to it on broader grounds,¹⁴ and thus reject such meat. Nevertheless, while the Advocate General’s Opinion squarely addressed public concern in this way, the Court itself chose to base the relevant part of its ruling solely on the narrow, initial point about the lack of a legal duty on the part of the Council to take the report into account.¹⁵

*Fedesa*¹⁶ concerned a challenge to the validity of the successor¹⁷ hormones Directive 88/146, which similarly prohibited the use of certain hormones in livestock farming. As the Court pointed out,¹⁸ the first recital in the preamble to the Directive stated that assessments of the effect on human health of those substances varied in the different national regulations. The second recital stated that the Council therefore considered it necessary to lay down rules ensuring that “all consumers are able to buy the products in question under largely identical conditions of supply and that these products correspond to their anxieties and expectations in the best possible manner”. The recital went on to state that the Council considered that such a course of action was bound to bring about an increase in consumption of the product in question. In essence, the case was centred around these first two recitals.

¹³ *ibid.*, emphasis added.

¹⁴ See A.G. Mischo in C-331/88 *R v MAFF, ex parte Fedesa* [1990] E.C.R. I-4023, where he observes that public campaigns against the use of hormones in agriculture were not “solely based on fears as to the harmful nature of hormones, but were also the manifestation of a more general trend in public opinion, namely a growing aversion on the part of the public to the use of chemical products in agriculture” (para.16).

¹⁵ Paras.35-36 of the Judgment.

¹⁶ n.14 above.

¹⁷ The earlier Dir. 85/649 having been annulled in *UK v Council*, n.7 above (not, as we have seen, for having the wrong legal base or for insufficiency of reasons, but on another, more technical procedural ground).

¹⁸ Para.6.

The applicants challenged the Directive on a number of grounds, but only those of legal certainty, legitimate expectations and proportionality involved the issue of public concern. In relation to the former two, it was argued during the proceedings “that the directive lacked any scientific basis justifying the public-health considerations and consumer anxieties which underlay its adoption and that it frustrated the legitimate expectations of traders, who were entitled to expect that the substances in question would not be prohibited in the absence of any objectively based doubt as to their safety, efficacy and quality.”¹⁹ On the initial issue of legal certainty *stricto sensu*, the Court was prepared to assume that the principle requires any measure adopted by the Community institutions to be founded on a rational and objective basis. It is, of course, science which tends to be regarded as rational and objective, and the Commission’s scientific evidence had been that the relevant hormones posed no health risk to consumers.²⁰ In these circumstances, one might think that the ban was not founded on a rational and objective basis, because science did not support it. Nevertheless, according to the Court, this was a highly discretionary policy area where the Court would only intervene if the Council had manifestly exceeded the limits of its discretion. In the Court’s view, it was not therefore necessary for the Council to prove that the Commission’s evidence as to the safety of the relevant hormones was unsound. Faced “with the divergent appraisals by the national authorities of the Member States . . . the Council remained within the limits of its discretionary power in deciding to adopt the solution of prohibiting the hormones in question, and respond in that way to the concerns expressed by the European Parliament, the Economic and Social Committee and by several consumer organizations.”²¹

At first sight, one might think that use of the rather technical word “appraisal”²² implies that the divergent Member State appraisals were scientifically based. On this view, the Court was stating that the Council had not committed a manifest error in not following the *Community’s* rational and objective science because its action was supported by the also rational and objective scientific appraisals of a number of *Member State* authorities.²³ However, if one looks at the UK’s submissions in the case, it becomes clear that no scientific evidence had been adduced in support of those Member

¹⁹ Para.7.

²⁰ See A.G. Mischo’s Opinion, para.15.

²¹ Para.9. On these appraisals and concerns, see further A.G. Mischo’s Opinion, para.16. Scott and Vos have drawn attention to the fact that the concerns taken into account go beyond those held by the directly elected European Parliament (EP), thus suggesting that the democratisation of risk goes beyond the confines of representative democracy alone (“The Juridification of Uncertainty: Observations on the Ambivalence of the Precautionary Principle within the EU and the WTO” in Joerges and Dehousse (eds.), *Good Governance in Europe’s Integrated Market* (OUP, 2002), p.253 at p.284).

²² Or indeed the word “assessment” used in the first recital of the preamble to the Dir., which perhaps suggests a scientific risk assessment.

²³ On this reading: the fact that the ban also reflected wider public concerns (as expressed by the EP, the ESC and consumer organisations) is presented as a secondary, supporting factor; and, unusually, public concerns here were backed up by Member State scientific support.

State appraisals or assessments.²⁴ In other words, it is more likely that the “divergent appraisals” were non-scientific and based on national public concern. Seen from this stance, the ruling can be seen as offering explicit support for the consideration of public concern over science.²⁵

Neither, according to the Court, did the Directive frustrate the legitimate expectations of traders. First, traders should not have assumed that the Community scientific advice would be followed since, under the terms of the relevant Directive, only the Commission and not the Council were legally obliged to take this advice into account.²⁶ And secondly, “in view of the divergent appraisals which had been made, traders were not entitled to expect that a prohibition on administering the substances in question to animals could be based on scientific data alone.”²⁷ Again, there are two possible interpretations of this latter statement. On the one hand, it could be interpreted to mean that, given that there were *competing* scientific views at play (national authorities versus Community), it was impossible for traders to expect the prohibition to be founded on just science and that public concern was thus bound to play a deciding role. On the other hand – and more plausibly given the UK’s submissions mentioned above – it could be seen as lending support to a reliance on broad EU public opinion *in place of* science.

As for proportionality, the applicants had argued that the Directive infringed the principle, *inter alia*, because an outright ban was not necessary: consumer anxieties could have been allayed using straightforward information and advice.²⁸ In a somewhat cryptic passage, the Court replied to this as follows:

“As regards the arguments which have been advanced in support of the claim that the prohibition in question is not necessary, those arguments are in fact based on the premise that the contested measure is inappropriate for attaining objectives other than that of allaying consumer anxieties which are said to be unfounded. Since the Council committed no manifest error in that respect, it was also entitled to take the view that, regard being had to the requirements of health protection, the removal of barriers to trade and distortions of competition could not be achieved by means of less onerous measures such as the dissemination of information to consumers and the labelling of meat.”²⁹

What the Court appears to be saying here is that the measure was potentially appropriate for securing a number of objectives. One of these objectives was

²⁴ See A.G. Mischo’s Opinion, para.10.

²⁵ Scott and Vos’ brief discussion of the case is along these lines. They are critical of the Court for taking into account broad public concern without fully examining the status of these concerns or questioning their legitimacy (n.21 above, pp.284-285).

²⁶ Para.10. *Cf. UK v Council* (n.7 above), where a similar point was made in relation to a claim about insufficiency of reasons (as opposed to legitimate expectations here).

²⁷ Para.10.

²⁸ Para.12.

²⁹ Para.16.

allaying (allegedly unfounded) consumer anxieties and the Council had committed no manifest error in concluding that a ban was necessary to achieve this objective. However, that was by no means the only one. Others included health protection and the removal of barriers to trade and distortions of competition and, according to the Court, the Council was again entitled to conclude that a ban was necessary to achieve these objectives and that lesser, informational measures would not be sufficient.³⁰

Having tackled hormones, the Community set its sights next on the use of antibiotics in agriculture. Again, its controls in this area gave rise to a number of challenges by disgruntled producers. *Boehringer*³¹ involved annulment proceedings brought against a Community ban on the marketing of beta-agonists for use in livestock production. As the Court of First Instance (CFI) noted, these antibiotics had two uses: one, in low doses, as a veterinary medicinal product which did not threaten public health; and the other, in extremely high doses, for the purpose of fattening cattle, which was potentially harmful to humans (producing sickness on eating the relevant meat).³² The ban covered both of these uses. *Boehringer* claimed that this *general* ban was disproportionate and that specific use for therapeutic purposes should have been allowed to continue, with appropriate control measures in place to prevent unscrupulous farmers from abusing the therapeutic option. The CFI disagreed. It noted first that proportionality involves considering, *inter alia*, whether the measures in dispute are necessary for achieving their aim.³³ The aim or objective of the relevant Directive³⁴ was said to be twofold: to protect public health and to meet consumer concerns as required by the common agricultural policy.³⁵ In relation to the public health aim, the Court stated that it could not be demonstrated that the Council had made a manifest error of assessment in holding that the general ban was preferable from the point of view of protecting public health, particularly given the uncertainties surrounding the effectiveness of control measures on therapeutic uses and their high cost to the public purse.³⁶ In relation to the common agricultural policy aim, the CFI ruled that:

“Furthermore, even if protection of public health takes priority over all other considerations, the protection of consumer confidence is equally important. It is undisputed that at least a section of the public and the professions involved together with many Members of the European Parliament were in favour of an outright ban on beta-agonists, and that their anxieties would not have been allayed by the establishment of

³⁰ In this respect, the Court’s judgment appears to mirror the A.G.’s Opinion (see paras.17, 29, 33, and 39-42 of his Opinion). In reality of course, these “alternative” objectives are also potentially associated with public concern – a point that will be returned to towards the end of this section.

³¹ Cases T-125/96 and T-152/96 [1999] E.C.R. II-3427.

³² *i.e.*, with this latter use, it is antibiotics rather than hormones being used as growth-promoters.

³³ Para.76.

³⁴ 96/22.

³⁵ Para.77.

³⁶ Paras.90-97.

control mechanisms, however effective they might have been in purely operational terms. It is also undisputed that, in many Community countries, information campaigns going so far as to issue calls for the boycott of meats containing hormones were launched by consumer associations in the past. In those circumstances, and taking into account the growing use of beta-agonists for the purpose of fattening cattle artificially, the Council cannot be regarded as having committed a manifest error of assessment in holding that only a general prohibition was capable of restoring consumer confidence.”³⁷

Perhaps not surprisingly given the dual objective of the Directive, the CFI here is unequivocal in attaching equal importance to consumer perceptions of risk. Even if specific control mechanisms were completely effective in tackling the risk to public health from potential abuse of the therapeutic availability of beta-agonists, public anxiety or fear, albeit scientifically unjustified, would still it seems have supported a general ban of both growth promoting and therapeutic uses.

The more recent *Pfizer*³⁸ case involved a challenge on the grounds of, *inter alia*, misuse of powers. At the heart of the case was an antibiotic called virginiamycin, which was added to animal feedingstuffs in low doses to act as a growth promoter. Although scientifically unproven, the risk was that the antibiotic resistance developed in animals to this drug would be transferred to humans.³⁹ This risk led certain Member States and eventually the Community itself to ban its use as a feed additive. The ban was implemented by way of a Regulation,⁴⁰ based on powers in the relevant Directive,⁴¹ which only provided for action to be taken on human or animal health, or environmental grounds. Pfizer alleged that the Community ban was actually driven by the sole aim of “creating a favourable political impression in the eyes of the press and public opinion, which is tantamount to a misuse of powers.”⁴² In response, the CFI ruled that this was not borne out by the documentary evidence, which revealed that the contested regulation pursued, above all else, public health objectives.⁴³ It then added that:

In any event, the restoration of consumer confidence can in such circumstances also be an important objective which may justify even substantial economic consequences for certain traders.⁴⁴

The case thus appears to suggest that the restoration of consumer confidence can be a legitimate secondary objective alongside the protection of public health even when – as in this case – the relevant legal base for the measure does not provide for any action on that basis and where the restoration of

³⁷ Paras.98-100.

³⁸ Case T-13/99 *Pfizer Animal Health SA v Council* [2002] E.C.R. II-3305.

³⁹ Paras.30-41.

⁴⁰ 2821/98.

⁴¹ Art.11(3) of Dir. 70/524, as amended.

⁴² Para.455. See also para.407.

⁴³ Para.462.

⁴⁴ *ibid.*

consumer confidence is not explicitly mentioned as an objective in the measure itself.

Just as the *Pfizer*, antibiotics case involved a misuse of powers challenge, so in the principal UK BSE⁴⁵ case,⁴⁶ the UK similarly claimed, *inter alia*, that the Commission's Decision 96/239 – which had imposed a ban on exports of live cattle and beef products from the UK – was unlawful on this ground. The UK maintained that the relevant Directives⁴⁷ on which the Commission's Decision was based provided only for measures to be taken on health grounds and that the Commission had in fact acted for economic reasons to reassure consumers, to stabilise the relevant markets and to safeguard the beef industry.⁴⁸ The relevant recitals of the preamble to the Decision were as follows:

“Whereas, under current circumstances, a definitive stance on the transmissibility of BSE to humans is not possible; whereas a risk of transmission cannot be excluded; *whereas the resulting uncertainty has created serious concern among consumers.*”⁴⁹

The export ban Decision was prompted by the statement of the UK Spongiform Encephalopathy Advisory Committee (SEAC), to the effect that exposure to BSE was the most likely explanation of the recent cases of variant Creutzfeldt-Jakob Disease (vCJD) in humans. Nevertheless, SEAC was also careful to point out that there was no direct evidence of a link, and the EU Scientific Veterinary Committee (SVC) endorsed this.⁵⁰ It was, in other words, still a matter of scientific uncertainty: scientists were not claiming to be certain of a link. Furthermore, neither SEAC, nor, initially, the SVC, had recommended going much beyond the existing controls⁵¹ on BSE, which had been in place for some time.⁵² While the UK Government initially followed these recommendations, it very quickly decided to go beyond the scientific advice it had been given to introduce the Over Thirty Months Scheme (OTMS), under which all of the highest-risk, older cattle would be kept out of the food chain.⁵³ The reasons for this were partly

⁴⁵ Bovine Spongiform Encephalopathy or “mad cow disease”. For a wide-ranging analysis of a number of the risk regulatory issues associated with the BSE crisis, see further Little, “BSE and the Regulation of Risk” (2001) 64 M.L.R. 730. For discussion of the EU aspects, see Neyer, “The Regulation of Risks and the Power of the People: Lessons from the BSE Crisis” (2000) 4(6) *European Integration Online Papers*; and Joerges (n.6 above). On comitology aspects, see Krapohl, “Risk Regulation in the EU Between Interests and Expertise: The Case of BSE” (2003) 10 J.E.P.P. 189.

⁴⁶ Case C-180/96 *UK v Commission* [1998] E.C.R. I-2265.

⁴⁷ Art.9 of Dir. 89/662, as amended; and Art.10(4) of Dir. 90/425, as amended.

⁴⁸ Para.35.

⁴⁹ Emphasis added.

⁵⁰ Case C-180/96 (n.46 above), paras.9 and 13.

⁵¹ Such as removing high-risk “Specified Bovine Offal” (SBO – e.g. head and spinal cord) from the food chain.

⁵² SEAC's principal recommendation was that cattle over 30 months be deboned in licensed plants, with the nervous and lymphatic tissue trimmings designated as SBO and thus kept out of the food chains.

⁵³ *i.e.* the whole carcass and not just SBO material.

practical,⁵⁴ but were partly also a matter of accepting the need to go beyond the science in order to reassure the public and maintain public confidence in British beef.⁵⁵ The major supermarkets had made it clear that they would not sell any meat from cattle aged over 30 months.⁵⁶ As in the case of GMOs – where many supermarkets have removed GM products from their shelves – the supermarkets here were effectively acting as a gauge of and also to preempt the consumer concern associated with beef. On this occasion, where the supermarkets led, the Government felt bound to follow.

Thus, we have seen that the OTMS was designed by the UK, in part, to restore consumer confidence in British beef and hence to stabilise the UK beef market. However, the UK was concerned that the Commission had inappropriately based its even more draconian additional measures (in the shape of an export ban of all cattle and beef products of whatever age) on exaggerated consumer concerns and a desire to stabilise the EU beef market rather than to protect public health.⁵⁷ This does not necessarily indicate double standards on the UK's part. One assumes that the UK was challenging the Commission's measures not because they were based on consumer confidence *per se* (which would have involved double standards given the basis of the OTMS), but because the Community measures were a disproportionate reaction to the confidence issue.⁵⁸ In any event – whatever its own stance happened to be – the UK was in a position to argue that the Commission's approach was unlawful.

In the event, the UK failed in its argument before the Court. In accordance with its standard case law, the Court ruled that, for a misuse of powers to be established, the Community institution must have adopted the measure with the exclusive or main purpose of achieving an objective other than that specifically provided for.⁵⁹ It stated that, whilst the objectives of the Commission's decision were to be sought in the recitals of the preamble, the recitals had to be considered as a whole rather than in isolation.⁶⁰ Read as a whole, the recitals showed that the Commission was prompted by public health concerns,⁶¹ and there was no evidence to suggest that the exclusive or main purpose was economic in nature.⁶² The UK National Farmers' Union (NFU) adopted much the same⁶³ argument in its indirect, national court

⁵⁴ There were capacity problems associated with the recommended deboning operations.

⁵⁵ *The BSE Inquiry: The Report* (Stationery Office, 2000), paras.7.381-7.382 and 7.387, available at <http://www.bseinquiry.gov.uk/report/index.htm>. The Report suggests that a third factor may have been a desire to make Europe look more favourably on British beef (para.7.388).

⁵⁶ *ibid.*, paras.7.382 and 7.387.

⁵⁷ Neyer (n.45 above) at p.5.

⁵⁸ I am grateful to Veerle Heyvaert for this observation.

⁵⁹ Para.64.

⁶⁰ Paras.65-66.

⁶¹ Para.66.

⁶² Para.67.

⁶³ Though not identical: the NFU's submission was based entirely on consumer concern, as opposed to the UK's argument which, as we have seen, twinned consumer concern with economic purposes of market stabilisation.

action challenging the export ban,⁶⁴ and – naturally enough in a judgment issued on the same day as that in *UK v Commission* – the Court effectively echoed the reasoning of its judgment in the direct action.⁶⁵

Finally, *Commission v Council*⁶⁶ involved a legal base challenge brought by the Commission to a Council Regulation on the traceability of cattle and the labelling of beef. The Commission contended that the Regulation was predominantly a public health measure and thus ought to have been based on Article 100a (now 95), requiring the involvement of the European Parliament (EP) via the co-decision procedure. However, the Council – with which the Court agreed – argued that the measure was more concerned with stabilising the market in beef and restoring consumer confidence via increased transparency. For that reason, the legal base adopted – Article 43 (now 37) relating to agriculture, with its lesser involvement for the EP – was entirely appropriate. The fact that the measure may also have had a public health element did not alter this conclusion, since taking into account public health in the context of measures adopted under Article 43 was entirely appropriate under the third subparagraph of Article 129(1) (now 152(1)) of the Treaty,⁶⁷ and public health was not the principal aim of the measure.

To summarise the position so far, in the *UK v Commission* (BSE), *NFU*, and *Pfizer* cases above, the applicants argued for the annulment of a measure on the explicit ground that it was principally and thus unlawfully aimed at restoring consumer confidence rather than public health. Similar, though more implicit, arguments relating to public concern were also presented in *Boehringer*, where the applicant claimed that the relevant measure was not necessary for the protection of public health, and in the reasons-related part of the UK's application in *UK v Council*, where it was argued that the health-basis of the measure was not scientifically justified. And in *Fedesa*, the applicant was arguing that neither the public health nor the consumer concern aims of the measure were scientifically justified, and that the measure was not necessary for allaying consumer concern. In all of these cases, in other words, the applicants were contesting the health-basis of the relevant measures. More often than not, the argument (explicit or implicit) was that they were more about tackling public concern than addressing real health issues. In contrast, in *Commission v Council*,⁶⁸ and in the legal base part of the claim in *UK v Council*, the applicants argued something like the reverse position – that the relevant measure was principally aimed at protecting public health rather than addressing consumer concern-related issues. However, in neither set of cases were the applicants' arguments

⁶⁴ Case C-157/96 *R v MAFF, ex parte National Farmers' Union* [1998] E.C.R. I-2211, para.42.

⁶⁵ See paras.44-46. Although, reflecting the slightly different submissions of the relevant parties noted at n.63 above, para.46 speaks of there being no evidence that the main purpose was to allay "consumer concern" (as opposed to the UK case, which speaks of "economic" purposes).

⁶⁶ Case C-269/97 [2000] E.C.R. I-2257.

⁶⁷ Which requires the integration of human health protection into all Community policies and activities. This Art., introduced at Maastricht, was not in existence at the time of *UK v Council* (n.7 above), though as we have seen, the Court there also ruled that health was integral to agricultural policy.

⁶⁸ n.66 above.

regarding consumer concern successful.⁶⁹ As we have seen, in the former cases, the Court either held that the relevant measures were predominantly aimed at protecting public health, or that they were appropriately and necessarily aimed at both public health and public concern issues. In the latter cases, it ruled either that health as an objective did not affect the choice of legal base, or that consumer concern and not health was the predominant objective.

What then can we conclude from the examination of the case law presented above? First, it is apparent that in virtually all of the cases in which public concern has been raised, the Court has produced a result upholding the relevant Community measure.⁷⁰ Secondly, it is clear from the cases that, where public concern has been directly raised in EU law, it has been in a *negative* manner. In other words, the argument is that the EU authorities *should not* have based the relevant legislative measure on public or consumer concern. There are no *positive* cases where public interest-type applicants have argued that the authorities ought to have based a measure on public concern but have not done so. In this respect, public concern is unlike the precautionary principle, where there are both cases involving arguments that the Community authorities have been excessively or overly precautionary (negative cases) *and* cases arguing that they have been insufficiently precautionary (positive cases).⁷¹

Of course, public concern is also unlike the precautionary principle in that there is no explicit mention of it in the Treaty as a principle which must be taken into account in decision-making. With the precautionary principle, in contrast, there is an explicit statement to this effect in Article 174 EC which has been mainstreamed⁷² as a general principle of law by the *Artegodan* case.⁷³ That said, it is worth recalling that public concern is often closely linked with notions of market stability⁷⁴ and there is of course a duty in the Treaty which requires attention to the creation of such stability.⁷⁵ Nevertheless, it ultimately depends on the *specific* enabling power or legal base for a relevant measure. If a legal power only allows action to be taken by the Community to protect public health or the environment, then one clearly cannot point to a duty elsewhere in the Treaty to stabilise the market

⁶⁹ Other than in *UK v Council* (n.7 above) where, as noted in n.17 above, they succeeded on an unrelated, procedural ground.

⁷⁰ With the exception of *UK v Council*, *ibid.*

⁷¹ Albeit that there are relatively fewer of the latter – Heyvaert, “What’s in a Name? Assessing the Impact of the Precautionary Principle in European Community Law and Policy”, paper presented at IALS, 9 November 2004 (who categorises the cases into “excessive precaution” and “insufficient precaution” challenges).

⁷² *I.e.* it has been applied beyond the strict confines of environmental policy in Art.174.

⁷³ Cases T-74/00, T-76/00, T-83/00, T-84/00, T-85/00, T-132/00, T-137/00 and T-141/00 *Artegodan and Others v Commission* [2002] E.C.R. II-4945, para.184.

⁷⁴ See the very start of this section, where it was suggested that consumer confidence (associated with market stability) may act as a proxy for public concern.

⁷⁵ Art.33(1)(c) EC within the Agricultural Title, which states that “The objectives of the common agricultural policy shall be . . . to stabilise markets”.

as a *sole or principal* justification for that particular measure.⁷⁶ Equally, if the relevant enabling power allows only action to stabilise the market, then one cannot introduce a measure based solely or principally on public health grounds. Of course, if a power allows for both – as we have seen in some of the case law above – then there is no problem: one can clearly have public health and consumer concern as relevant objectives and either can be given predominant status.

That said, the above assumes a degree of honesty as to legislative purpose. In practice, however, public concern can effectively be *disguised*. Thus, if a legal power only allows for action on health or environmental grounds and one is being honest, then public concern may only form a minor purpose. However, in practice, there is nothing to stop the Community authorities from disingenuously claiming that a measure is solely or predominantly health/environment-based when it is actually predominantly motivated by public concern. Given the low intensity style of review adopted by the Court in highly discretionary scientific areas, applicants will face an uphill struggle in persuading the Court to intervene in such cases. The only real barrier is the *Pfizer* case, which requires there to be at least some scientific evidence behind a health/environment measure.⁷⁷ However, while this means that the authorities cannot dress up a public concern measure in the clothing of health/environment if there is no scientific foundation whatsoever, it does not prevent lesser but still significant examples of purpose non-transparency.

Of course, the preceding discussion also presupposes that public health and public concern are separate matters.⁷⁸ If, in contrast, one adopts a less rigid distinction or separation between the two, then it no longer appears disingenuous to claim that, in public concern cases, one is acting on health grounds. Ultimately, it perhaps comes down to the essential fluidity and contestability of the notion of acting on public health grounds. I have taken acting on public health grounds to mean that there are established scientific reasons for doing so.⁷⁹ From this perspective, the distinction drawn between public health and public concern is a real one – public health is scientifically backed and public concern is not. However, it is also perfectly possible to argue that an institution is still, in a sense, acting on public health grounds where there is little or nothing in the way of scientific justification but where the public is nevertheless concerned for its health. From this latter perspective, the distinction drawn between public health and public concern (in a narrow, health risk perception sense) is a tenuous or artificial one. With the MMR vaccine or mobile phone masts for example – if an authority supports the availability of single vaccines or denies planning permission for a mast, can it be said to be acting on public health grounds or solely on the basis of public concern? One's answer will depend not only on one's view of the relevant science but, more importantly, on what one's conception of acting on public health grounds involves.

⁷⁶ Though if it remains a *minor or subsidiary* objective, then the *Pfizer* and *UK v Commission* (BSE) cases (nn.38 and 46 above, respectively) suggest that there will be no illegality.

⁷⁷ See the discussion of *Pfizer* in n.135 below.

⁷⁸ *Cf.*, similarly, public concern and science – n.4 above.

⁷⁹ *ibid.*

Public Concern, the EU and the WTO

Having examined the place that public concerns surrounding risk activities play in EU case law, it is instructive to see how this compares with the treatment of public concern in WTO jurisprudence. Like the previous section, this one is concerned only with how the issue has been dealt with in the cases themselves (all of which involve the EU). It is beyond the scope of this article to attempt to analyse public concern within EU law or WTO law more generally.⁸⁰

The Beef or *Meat Hormones* case is central to any such comparison. In the internal context, we have already examined Court of Justice case-law which suggests that the relevant EC Directive banning hormones in beef was driven by consumer anxieties and broader concerns rather than scientifically-grounded safety issues. Not surprisingly, in the external context, similar accusations were levelled at the EU in WTO dispute settlement proceedings. Before the Panel,⁸¹ Canada stated that the EU measures had been motivated by four sets of concerns: (i) anxiety concerning the danger to human health caused by illegal use of growth promoters; (ii) the pressure of public opinion; (iii) the economic consequences of a “sensationalist campaign”, which had produced a collapse in the veal market and a sharp decline in beef sales across the EU; and (iv) distortions in the conditions of competition among EU Member States due to differing regulations on growth hormones.⁸²

New Zealand claimed that Article 5 of the SPS Agreement required sanitary or phytosanitary measures which provided a higher level of protection than an international standard,⁸³ to be justified by way of a scientific risk assessment. According to New Zealand, no such risk assessment had been carried out in relation to the six hormones which formed the subject of the dispute. While, in its view, “it was evident that consumers in the European Communities might harbour concerns about the use of these substances, consumer perceptions as to risk did not provide a basis for the adoption of measures under Article 5.”⁸⁴ The US similarly argued that a risk assessment was required because, to be legitimate, a sanitary measure must be seeking to protect the public from a specified risk. According to the US, “the remarkable characteristic of the public debate in the European Communities on these hormones was that the “risk” was usually described in terms of consumer anxieties rather than any observable adverse effect on human health. During the consultations, the European Communities had failed to

⁸⁰ *E.g.* analysis of public concern issues associated with EU policy on GMOs, on which see Scott (n.3 above).

⁸¹ There are two Panel Reports – one relating to the complaint brought by Canada and the other by the US. Though very similar in content, the two Reports will be considered separately here because there are interesting differences between them in relation to consumer concern. Both are available at <http://trade-info.cec.eu.int/wtodispute>

⁸² WT/DS48 - Measures affecting livestock and meat (Hormones) – henceforth “Canada”, para.IV.6. See also para.IV.10.

⁸³ The relevant international standard here being the one developed by the Codex Alimentarius Commission in relation to growth-hormones.

⁸⁴ WT/DS26 - Measures affecting meat and meat products (Hormones) – henceforth “US”, para.V.28.

identify any specific risk to human or animal life or health against which the ban was designed to protect.”⁸⁵

In fairness, however, it does not appear that the EU was claiming that consumer perceptions should have a place within a risk *assessment* within Article 5, since the US Panel tells us that “none of the parties has argued that factors not listed in Article 5.2, such as consumer preferences, can be taken into account in a risk assessment in accordance with Article 5.”⁸⁶ The EU was, however, contesting the claim that the purpose of its measures was to allay consumer anxieties rather than to protect human health. At one point in the Panel’s report, we are told that the EU “asserted that the historical record clearly demonstrated that the purpose of the measures was to protect human health and animal health from risks arising from the use of the hormones at issue.”⁸⁷ However, elsewhere in the Panel report, it seems that the EU was in fact denying only that “consumer anxieties” were the *sole* purpose for which the measures were adopted.⁸⁸ Moreover, the EU was at pains to distinguish between risk assessment, which provides the scientific basis for regulatory action, and risk management, which is the process by which elected representatives decide what action to take in the face of the risk assessment presented to them by the scientists.⁸⁹ According to the EU, such action is likely to be based on a range of factors, including public health, environmental protection, social, economic and political values and consumer concerns.⁹⁰ In the context of the measures on hormones, the EU thus admitted that it was likely that consumer concerns had been taken into account during the risk *management* phase, since consumer concerns on potential risks to human health resulting from the use of hormones were very high at that time (and were even higher today).⁹¹ The EU also claimed that it was not alone in doing so – pointing to the way in which consumer concerns had clearly influenced US policy.⁹²

Although the Panel could be said to have implicitly accepted this particular argument relating to consumer concern,⁹³ it came back to the issue in relation

⁸⁵ US, para.IV.110.

⁸⁶ US, para.8.105. The impact of the subsequent Appellate Body (AB) ruling on this particular issue will be considered at the end of this section.

⁸⁷ Canada, para.IV.14. See also para.IV.18 and US, para.IV.21.

⁸⁸ US, para.IV.113.

⁸⁹ US, para.IV.97.

⁹⁰ *ibid.* Similar claims have been made by the Commission before the Court of Justice, though the Court itself has never explicitly responded to these (see *e.g.* Cases T-344/00 and T-345/00 *CEVA v Commission* [2003] E.C.R. II-229, para.66; and A.G. Alber’s Opinion in Case C-248/99 P *France v Monsanto and Commission* [2002] E.C.R. I-1, para.120).

⁹¹ US, para.IV.113 and see also para.IV.14.

⁹² US, para.IV.113 (“There was little question that consumer concerns about safety and other issues also influenced the US agencies’ decisions, even though there was arguably little, if any, basis for those concerns”) and IV.124 (citing the US Delaney Clause, relating to additives, which was enacted in the context of public fear of cancer).

⁹³ US, para.8.160. *Cf.* the subsequent AB ruling querying the Panel’s distinction between risk assessment and risk management, discussed further at the end of this section.

to Article 5.5 of the SPS agreement.⁹⁴ Article 5.5 involves three distinct elements, all of which must be demonstrated if a breach of that Article is to be found. First, a Member must have established for itself different levels of sanitary protection in different situations. Secondly, the differences in the levels of protection must be arbitrary or unjustifiable. And finally, those arbitrary or unjustifiable differences must result in discrimination or a disguised restriction of international trade.⁹⁵ For present purposes, we are concerned only with the last of these three elements because it was only in relation to that element that the issue of public concern was raised. The question of whether there had been discrimination or a disguised restriction of international trade was itself raised in relation to three separate factual differences in levels of protection. First, there was the difference in levels of protection between residues of the three *natural* hormones in dispute when used as growth promoters and residues of the same natural hormones present endogenously⁹⁶ in meat and other foods.⁹⁷ Secondly, there was the difference in levels of protection between the relevant *synthetic* growth promoting hormones in dispute and natural hormones in dispute present endogenously.⁹⁸ And thirdly, there was the difference between the relevant growth promoting hormones in dispute and carbadox and olaquinox (antimicrobial growth promoters used as feed additives in pig farming).⁹⁹ In answering the question, in all three cases, of whether there had been discrimination or a disguised restriction of international trade within Article 5.5, the Panel drew attention, *inter alia*, to the objectives other than human health that the EU had had in mind when enacting or maintaining its ban on the use of natural hormones for growth promotion purposes.¹⁰⁰ By these “other” objectives, the Panel pointed to the EU’s objective of harmonising the differing EU Member State schemes, removing barriers to intra-Community trade in beef, increasing the consumption of beef, and reducing beef surpluses.¹⁰¹ According to the Panel such action provided more favourable treatment to domestic EU producers. While consumer concern is not explicitly mentioned in the main text here,¹⁰² it clearly forms the underlying initial impetus to harmonisation: without differing consumer concerns, Member State regulations would not have been so different. Furthermore, when we come, later in the article, to the Appellate Body’s (AB) response to this conclusion by the Panel, it will be seen that the AB examined the issue through much more of a consumer concern lens.

Finally, apparently in order to avoid any misunderstanding as to the scope and implications of the findings in its Report, the Panel was keen to stress, *inter alia*, that it “had not addressed the ability of any Member to enact

⁹⁴ Which provides that, in applying their chosen level of sanitary or phytosanitary protection, Members “shall avoid arbitrary or unjustifiable distinctions in the levels it considers to be appropriate in different situations, if such distinctions result in discrimination or a disguised restriction on international trade.”

⁹⁵ Canada, para.8.177.

⁹⁶ *i.e.* occurring naturally.

⁹⁷ Canada, para.8.204.

⁹⁸ Canada, para.8.210.

⁹⁹ Canada, para.8.222.

¹⁰⁰ Canada, paras.8.207, 8.220, and 8.245.

¹⁰¹ Canada, para.8.207.

¹⁰² Though it does appear in Canada, footnote 462.

measures which are intended to protect not consumer health but other consumer concerns.”¹⁰³ While the above analysis of consumer concern might be said to give lie to such a claim, it seems that what the Panel had in mind here were regulatory measures such as labelling:

“we are aware that in some countries where the use of growth promoting hormones is permitted in beef production, voluntary labelling schemes operate whereby beef from animals which have not received such treatment may be so labelled.”¹⁰⁴

Labelling provides a less trade restrictive alternative to import bans.¹⁰⁵ Instead of public regulators taking into account consumer concerns in deciding to ban imports, imports are allowed, but labelling allows those consumers who perceive there to be a risk or who object on wider grounds, to avoid the relevant product.¹⁰⁶ This issue has particular resonance in the context of the current EU policy regarding GMOs,¹⁰⁷ where labelling has been chosen in place of banning imports – although labelling there is mandatory, whereas the Panel above mentions only voluntary schemes. While the labelling element does not form a part of the current dispute relating to the EU’s GMO regime,¹⁰⁸ it seems probable that it will form the basis of a subsequent complaint. It is not likely to be too long therefore before risk labels are tested for their WTO compatibility.¹⁰⁹

On appeal, public concern came to the fore in relation to the question of discrimination or a disguised restriction on international trade within Article 5.5. In this regard, the AB concerned itself only with the difference in levels of protection between the relevant hormones in dispute and carbadox and olaquinox.¹¹⁰ In reversing the Panel’s finding that this difference resulted in a disguised restriction on international trade, the AB appeared to accept that it was legitimate for the EU, with its mandate to establish a common market in beef, to act to address consumer concerns:

¹⁰³ Canada, para.8.278.

¹⁰⁴ *ibid.*

¹⁰⁵ On informational approaches to controlling traded product risks, see further Hilson, “Information Disclosure and the Regulation of Traded Product Risks” (2005) 17 J.E.L., forthcoming.

¹⁰⁶ Though of course it does not stop others from consuming, which may mean that the relevant wider concerns are not met. For those purely concerned on perceived health risk grounds, labelling is, in contrast, an effective solution.

¹⁰⁷ See Scott (n.3 above), pp.24-25.

¹⁰⁸ WT/DS291-WT/DS293 - Measures affecting the approval and marketing of biotech products (GMOs). Complaints have been brought by The US, Canada and Argentina regarding the alleged EU moratorium on licensing authorisations.

¹⁰⁹ On the WTO compatibility of the EU GM labelling regime, see *e.g.* Morgan and Goh, “Genetically Modified Food Labelling and the WTO Agreements” (2004) 13 R.E.C.I.E.L. 306; Makuch, “TBT or not TBT, That is the Question: The International Trade Law Implications of European Community GM Traceability and Labelling Legislation” (2004) 13 Eur. Env. L. Rev. 226; Hilson and French, “Regulating GM Products in the EU: Risk, Precaution and International Trade” in Cardwell, Grossman and Rodgers (eds.), *Agriculture and International Trade – Law, Policy and the WTO* (CABI, 2003), p.215.

¹¹⁰ Report of the Appellate Body, WT/DS26/AB/R and WT/DS48/AB/R (henceforth AB Report), para.241.

“We do not attribute the same importance as the Panel to the supposed multiple objectives of the European Communities in enacting the EC Directives . . . The documentation . . . makes clear the depth and extent of the anxieties experienced within the European Communities concerning the results of the general scientific studies (showing the carcinogenicity of hormones), the dangers of abuse (highlighted by scandals relating to black-marketing and smuggling of prohibited veterinary drugs in the European Communities) of hormones and other substances used for growth promotion and the intense concern of consumers within the European Communities over the quality and drug-free character of the meat available in its internal market. A major problem addressed in the legislative process of the European Communities related to the differences in the internal regulations of various Member States of the European Union (four or five of which permitted, while the rest prohibited, the use for growth promotion of certain hormones), the resulting distortions in competitive conditions and the existence of barriers to intra-community trade. The necessity for harmonizing the internal regulations of its Member States was a consequence of the European Communities’ mandate to establish a common (internal) market in beef.”¹¹¹

What is noteworthy is that while the EU, in its submissions, had claimed that public concerns were relevant to the second element of Article 5.5 (the arbitrary or unjustifiable element),¹¹² in relation to the third, disguised restriction element, it had in fact been trying to stress the predominant health motivation for its legislation.¹¹³ Ironically, the AB was, as we have seen, not at all bothered by the multiple objectives and, in clearing the EU of a breach here, in fact placed most emphasis on the ones based on consumer concern.

Another significant area where public concern has been at issue in WTO disputes is in relation to the concept of “like products” within Article III: 4 of the GATT 1994.¹¹⁴ Article III: 4 states that:

“The products of the territory of any Member imported into the territory of any other Member shall be accorded treatment no less favourable than that accorded to *like products* of national origin in respect of all laws, regulations and requirements affecting their internal sale, offering for sale, purchase, transportation, distribution or use.”¹¹⁵

In the *Asbestos* case, the Panel initially found that chrysotile asbestos fibres were “like” PVA, cellulose and glass fibres and, subsequently, that cement-based products containing chrysotile asbestos fibres were “like” cement-

¹¹¹ AB Report, para.245.

¹¹² AB Report, para.33.

¹¹³ AB Report, para.244.

¹¹⁴ Scott (n.3 above), p.24.

¹¹⁵ Emphasis added.

based products containing PVA, cellulose and glass fibres.¹¹⁶ On appeal, the AB reversed these findings. Like the Panel, it employed what has become a standard approach for analysing likeness, using four general criteria: the properties, nature and quality of the products; the end-uses of the products; consumer tastes and habits, or consumer perceptions and behaviour in respect of the products; and the tariff classification of the products.¹¹⁷ For present purposes, it is of course the third criterion, consumer tastes and habits, that is of most interest. The Panel had declined to take a position on this third criterion, on the basis that it would not provide clear results because tastes and habits are “very varied”.¹¹⁸ However, the AB stated that a panel could not decline to look into relevant evidence just because it may not be clear.¹¹⁹ In any event, the AB seemed to think that consumer tastes and habits were quite clear here. According to the AB, there was a proven health risk (of cancer) associated with chrysotile asbestos fibres.¹²⁰ Consumer perceptions of the risk associated with the different fibres and their associated products were therefore, inevitably, different; and these perceptions and the purchasing behaviour associated with them lent weight to the view that the fibres and their associated products were not “like”.¹²¹

The issue of like products within Article III: 4 was also raised in the *Meat Hormones* case. Here, the EU claimed that meat from animals treated with hormones was perceived by consumers as a distinct product and offered this as one of the reasons why growth hormone meat and ordinary meat could not be considered as like products.¹²² Canada argued that the perception of European consumers should not be relevant in determining the likeness of a product. In Canada’s opinion:

“The public authorities of WTO Members had a responsibility to educate the public and to make them aware of scientific facts. If the Panel were to allow “public perception” to become a factor in a “like product” determination, it would open the door to misapprehensions concerning scientific facts becoming the basis of a justification in the WTO for the adoption of discriminatory measures. Such an interpretation would obviously remove any incentive whatsoever for public authorities in the European Communities to make their populations aware of the scientific facts in respect of the six hormones at issue.”¹²³

In the event, the Panel and AB decided that the dispute fell to be decided by the SPS Agreement and not the GATT and thus the submissions on Article III: 4 were not addressed. This left open the crucial question of whether

¹¹⁶ WT/DS135 – Measures affecting the prohibition of asbestos and asbestos products, paras.8.144 and 8.150.

¹¹⁷ WT/DS135/AB/R, para.101.

¹¹⁸ Panel Report, para.8.139.

¹¹⁹ Para.120.

¹²⁰ Para.122.

¹²¹ *ibid.*

¹²² Canada, para.IV.338; US, para.IV.254.

¹²³ Canada, para.IV.348. Canada has put forward a similar claim in the current GMOs dispute – see n.108 above, Canada, First Written Submission, paras.317-318, available at http://www.genewatch.org/WTO/WTO_Submissions.htm

consumer perceptions of risk are relevant to the determination of product likeness only when – as in the *Asbestos* case – the health risks are scientifically proven. Alternatively – as in *Meat Hormones* – can consumer perceptions of risk also be taken into account where no scientific proof of harm has been firmly established? The GMOs dispute should cast some much needed light on this, as well as on the question of whether consumer perceptions as to process characteristics count in the same way as those associated with the physical character of the product itself.¹²⁴

Finally, it is worth noting the possible implications which the AB's ruling on the scope of risk assessment has for the role of public concern. The Panel, reflecting the submissions of the parties,¹²⁵ drew a distinction between risk assessment and risk management, stating that the former was scientific and the latter a non-scientific, policy exercise involving social value judgments made by political bodies.¹²⁶ The AB took issue with this apparently rigid distinction – which the Panel had employed to support a restrictive notion of risk assessment – holding that it had no basis in the text of the SPS Agreement.¹²⁷ What the AB was keen to allow as part of risk assessment was the human failings in terms of, for example, regulatory enforcement, which may have a key bearing on risks as they are found in the real world, as opposed to the laboratory.¹²⁸ However, it is rather a large jump from saying that risk assessment need not be confined to the laboratory-based, physical sciences to a claim that it should also encompass public concerns or anxieties in a more general sense. Public concerns may be relevant in a risk assessment if they are able to reveal, for example, particular frames or risk pathways which the experts may have missed.¹²⁹ Nevertheless, more general public anxiety about remaining uncertainties is probably not a legitimate factor for risk assessment under the SPS Agreement.¹³⁰

Conclusion

A comparison of the EU and WTO case law analysed above reveals some interesting points on which to conclude. What we find is that the EU, as an internal actor, is rather more constrained than it is as an external actor in its formal ability to take account of public concerns. Internally, the EU is constrained by the terms of its legislation in the various risk fields.¹³¹ If the relevant power allows only for action on health or environmental grounds, then the EU cannot rely on public concern – at least not formally and

¹²⁴ *ibid.*

¹²⁵ See *e.g.* n.89 above.

¹²⁶ See n.93 above.

¹²⁷ Which does not use the phrase risk management. AB Report, paras.181 and 205-6.

¹²⁸ AB Report, para.187.

¹²⁹ Winickoff *et al.*, n.2 above; *cf.* also n.4 above.

¹³⁰ As the AB states in its Report in *Meat Hormones* (at para.186): “. . . the Panel opposes a requirement of an ‘identifiable risk’ to the uncertainty that theoretically always remains since science can *never* provide *absolute* certainty that a given substance will not *ever* have adverse health effects. We agree with the Panel that this theoretical uncertainty is not the kind of risk which, under Article 5.1, is to be assessed.” See also US Panel, para.8.153, and n.4 above. It is also hard to see other *e.g.* ethical/political public concerns being held to be relevant.

¹³¹ To similar effect, see Scott (n.3 above) p.22.

explicitly – as a sole or predominant objective. This is true at both the risk assessment and risk management phases.¹³²

The AB ruling in the *Meat Hormones* case makes it clear, however, that the EU is much more free in terms of at least one of the WTO disciplines. Under the SPS Agreement, there must be an initial risk assessment on which the relevant measure is based, and public concern¹³³ cannot be considered as part of that scientific assessment. However, thereafter, the EU is formally and explicitly able to take into account public concern at the risk management stage.¹³⁴

In many ways, this is an ironic outcome, with WTO law seemingly giving more ground for the formal and explicit consideration of public concern than EU law. After all, it is not often that the WTO is considered to be more supportive of democracy than the EU. Of course, one might respond that to consider only the formal position is misplaced: after all, the EU can, as we have seen, surreptitiously act explicitly on scientific, public health or environmental grounds but as a cloak for public concern. And in doing so, it may be supported by a *substantive* version of the precautionary principle.¹³⁵ However, I would argue that public concern should be recognised as a legitimate ground for action in its own right rather than being hidden in this manner.¹³⁶ It makes no sense for the EU to argue one thing externally before the WTO and not be in a position to own up to it internally. There are two options. The first would be for the EU to examine its various, individual legal powers in the Treaty and elsewhere to see where public concern could be introduced as a relevant formal and explicit objective or consideration. As things stand, the existing market stability consideration within the Agriculture Title plays a somewhat limited role and, in any event, provides a rather indirect expression for public concern. More promising is, for

¹³² We are back, here, to the scope of public concern discussed in n.4 above. Scientifically justified public concerns can almost certainly be considered at both phases.

¹³³ Again, unless scientifically justified.

¹³⁴ Although, as we have seen, the AB cast doubt on the Panel's distinction between risk assessment and risk management, this was because of its restrictive effect on the scope of risk assessment. In principle, it remains useful to conceive of the post risk assessment elements of the SPS agreement (in particular Arts.5.4-5.6) as issues of risk management. The issue may be revisited in the forthcoming GMOs case (n.108 above – see the EU's Second Written Submission, paras.21-28, available at <http://trade-info.cec.eu.int/wtodispute>).

¹³⁵ *i.e.* the principle that one should be prepared to take action to protect human health or the environment even where there is scientific uncertainty about likely harm. Nevertheless, the *Pfizer* case makes it clear that the precautionary principle does not allow for measures based on hypothetical risks – there must be at least some scientific evidence of a risk (n.38 above, paras.140-146); *cf.* A.G. Mischo's Opinion in *Fedesa*, n.14 above, at para.17, allowing “purely hypothetical” risks to be tackled, which must now be considered wrong.

¹³⁶ To similar effect see Lee, *EU Environmental Law* (Hart Publishing, 2005) ch.9 (and also Lee, “Agricultural Biotechnology in the UK: Expanding the Debate?” in Everson, Scott & Vos (eds), *Uncertain Risks Regulated*, forthcoming), who argues that although public concern may have motivated a regulatory decision, it is often hidden away, with science made out to be the real justification. Lee, quite rightly, argues that this is harmful in terms of transparency and accountability.

example, the GM Food and Feed Regulation, which permits the Commission, in issuing its draft authorisation risk management decision, to take into account “other legitimate factors” beyond just the scientific risk assessment.¹³⁷

Alternatively, there could be a greater re-orientation of the precautionary principle along *procedural*, participatory lines,¹³⁸ so that the existing legal duty to take into account the precautionary principle itself takes care of the issue of public concern.¹³⁹ The advantage of this second option is that no Treaty or legislative amendments would be needed: the EU would be acting on existing health or environmental grounds/bases within the Treaty, but public concern would be seen as part and parcel of these via the precautionary principle. My own preference is for the former option: the danger with the latter is that it would further muddy the already murky waters of the precautionary principle, and taint what scientific respectability and mileage there is in the substantive version with what many will – however wrong-headedly – regard as the uninformed prejudices of populism. Whichever solution is adopted, considerable thought will have to be given as to how public concern is gauged and what pre-conditions are necessary for it legitimately to be taken into account. As Scott rightly notes, the last thing we want in the EU is the “uncritical, aggregation of ill-informed public preferences”¹⁴⁰ and some form of civic republican-type testing of public opinion will be needed in order to avoid this.¹⁴¹

¹³⁷ See Recital 32 and Art.7(1) of Reg. 1829/2003 on Genetically Modified Food and Feed [2003] O.J. L268/1. See further Lee, n.136 above.

¹³⁸ See e.g. Fisher, “Is the Precautionary Principle Justiciable?” (2001) 13 J.E.L. 315 at p.320.

¹³⁹ To similar effect, see Scott and Vos (n.21 above) p.279. Existing EU policy documents on the precautionary principle already make the link with public concern and stress the need for public involvement in risk regulation (see Council Resolution On The Precautionary Principle, Annex II to the Nice Presidency Conclusions 7, 8 and 9 December 2000, available at http://ue.eu.int/ueDocs/cms_Data/docs/pressData/en/ec/00400-r1.%20ann.en0.htm; and the Commission Communication on the Precautionary Principle, COM(2000) 1, paras.1 and 6.2 and para.5 of the summary).

¹⁴⁰ Scott (n.3 above) at p.26.

¹⁴¹ Scott makes some proposals of this type – *ibid.*, pp.28-29. See also Scott and Vos (n.21 above) pp.284-286.

BELFAST LEASES, LORD DONEGALL, AND THE INCUMBERED ESTATES ACT, 1849*

C.E.B. Brett

Introduction

Lawyers, and in particular solicitors, have in the past been notoriously reluctant to part with or destroy deeds, documents, papers and files which 'might one day come in useful'. For this reason, they have been for centuries invaluable as preservers of archives, especially title-deeds, wills, probates, and court pleadings – more so as most such documents had for centuries been laboriously written out in ink by scribes on parchment or durable paper. Today, photocopiers, computers, faxes, and emails have generated such enormous problems of storage that practitioners now shred the majority of documents every few years.

When, in 1953, as a young solicitor aged twenty-five, I was taken into partnership in the family firm of L'Estrange & Brett by my father and grandfather, problems of storage had already reached uncomfortable proportions. I was the sixth generation; the firm had been founded (as Ramsay & Garrett) in the last years of the eighteenth century. The attics and roof-space (not to mention the narrow but capacious strong-room) of the offices at no. 9 Chichester Street, Belfast, which the firm had occupied since 1886, were packed with piles and boxes of old documents, leaving no room for more. I was told that there had been a mild tidy-up in 1917, for fear of zeppelin raids; and another in 1939, for fear of air raids; but nothing else.

To make matters worse still, next door, the basement of no. 7 Chichester Street, was stacked from floor to ceiling with the unsorted bundles of paper and parchment that had come from the Shaftesbury estates¹ office, and the offices of Torrens & Bristow. These had accompanied old John Bristow, the last surviving partner in that firm, when he amalgamated his practice with that of L'Estrange & Brett in 1943 in order to provide himself, and his two spinster sisters, with an annuity in lieu of a pension. He and his redoubtable secretary Miss Nina Murdoch, with their humble clerk and scrivener Robert Meek (who laboriously inscribed in long-hand on parchment deeds of enduring importance, a task inconceivable today) occupied the ground floor of no. 7, and there looked after the affairs of sundry landed gentry and other profitable clients, including Lord Shaftesbury and Lord Harbington. This involved enforcing the covenants in their leases and fee farm grants, collecting their ground rents (then, before inflation eroded their value, a very lucrative commodity), their leets and heriots, and the payments for fishery rights, shooting rights, mineral rights, quarries, and such other ancillary

* This article is a slightly edited extract from Sir Charles Brett's book, "Georgian Belfast", published in 2004, and is here reproduced by permission of the Royal Irish Academy, Dublin.

¹ In 1857, Harriet, daughter of the third marquess of Donegall, married Anthony, eighth earl of Shaftesbury; on her father's death in 1883, his estates had mostly passed to her and so the Donegall estates became the Shaftesbury estates.

rights and royalties as they could collect. In short, they acted more like Scottish factors than modern Irish solicitors.

By 1960, the situation had become desperate. So, with the reluctant approbation of my father and grandfather, I opened negotiations on the firm's behalf with young Mr Kenneth Darwin, the energetic and efficient individual recently appointed to the office of Deputy Keeper of the Public Record Office of Northern Ireland. He was very anxious to get his toe into the doorway of the profession; I was equally anxious to find a solution to our own problems. In the end, after much hesitation and dubiety on the part of my partners, we reached a deal: the Public Record Office would take into its care and custody all the records and documents that were overwhelming us. It would take away, sort, catalogue, and preserve them in their lorry-loads; it would provide us with calendars, or catalogues, of all that had been entrusted to it. The documents were to be regarded as on permanent deposit, subject to the two over-riding provisos, that if any client of the firm wanted his documents back or if the firm for any reason wanted any of the documents, they should be immediately returned. Also, copies of documents deposited were to be freely available. Over a lengthy period, the documents so deposited were sorted out and catalogued, very competently, I believe mostly by the youthful Mr Brian Trainor. The Public Record Office was then still housed, conveniently for me, in the law courts at the foot of Chichester Street, where I could easily go and make searches during my lunch-hours.

This arrangement has worked, so far as I am concerned, to perfection. There were one or two worrying moments, as when I discovered that, inadvertently, I had consigned to the Public Record Office a parcel of letters describing in gynaecological detail an affair with a dancing-girl in Budapest, before the first war, of a client still living. These, and a few other sensitive bundles of early divorce papers, were quickly retrieved. But, so far as I know, no client has ever objected to the arrangement, or demanded the return of his papers, and it has provided a model for arrangements with a number of other firms. In his report for 1960–65 the Deputy Keeper of the records commented 'A quite outstanding collection of solicitors' records was deposited by Messrs L'Estrange & Brett of Belfast . . . over 150,000 documents of all kinds have been transferred from the firm's store-rooms . . . the variety is staggering . . . the value of such transfers of complete archives lies not merely in the large number and quality of the individual documents deposited but in the comprehensive nature of the whole collection . . . this collection illustrates, as no amount of comment can, why such collections should be preserved and deposited in full'.²

It was not long, however, until I discovered that I had, in a sense, shot myself in the foot. By presenting all these documents to the Public Record Office I had deprived myself of personal and immediate access to numerous papers, which could be helpful in my researches into the subject-matter of the book, ultimately published in 1967, upon which I soon afterwards embarked, *Buildings of Belfast, 1700–1914*.

² *Report of the deputy keeper of records, 1960–65* (Belfast, 1968), pp.9–10.

My first step had been to perambulate, slowly, every street in central Belfast, pencil and notebook in hand. I still have, and treasure, that notebook; I see that I started work on 1 January 1959. It records my, then rather amateurish, impressions of a great many buildings that no longer exist – over the past fifty years, fallen to decay, or to wicked bombers, or to grasping developers – so that the character of the city is utterly changed. I could then note that at no. 56 Chichester Street, parlour of a veterinary surgeon, there lived still a goose, two guinea-fowl, and many hens; Forde's cattle yard at no. 13 Eliza Street a 'gate opens to disclose two long low sheds, whitewashed, v. picturesque and farmyardy, decorated with stuffed sheeps' and cows' heads'; no. 20 McAuley Street a '1830ish pleasant horse-dealer's yard tunnelled through and behind two storey house'; East Street, no number, a 'defunct poulterer's shop, D. Gilmour & Son, with attractive duck and hen in plasterwork (?) or carved wood (?) on signboard'; and many other, long since gone, rural remains. Indeed, I remember, though I did not record, meeting a nice old lady sharing a stable with a donkey up an entry just off Cromac Street. Alas, it is not only the rustic remains that have disappeared: the late Georgian, Regency, early Victorian, high Victorian, and Edwardian buildings that still lent Belfast its highly individual character in the 1960s have mostly gone too.

As I completed my peregrinations, I started the work of research. The subject was a completely novel one, and I had more or less to start from the beginning. My principal aims were to track down the date, and architect (if any) or builder, of every building then standing. Fortunately, I had inherited my great-grandfather's fine library of Irish books, and of course had also close at hand the excellent Linen Hall Library. I procured photocopies of every known map of Belfast, had them framed, and hung them in sequence on the walls of our office staircase. The only complete set in Northern Ireland of that excellent journal the *Irish Builder* had just, maddeningly, been destroyed as junk by the egregious and bad-tempered Captain Young of Young & Mackenzie, architects. Fortunately, Ken Darwin was a man who knew when and how to break all the rules, and on his regular visits to Dublin as a trustee of the National Library borrowed and brought me back half a dozen volumes at a time. I went through every volume up to 1914 meticulously, card-indexing every Belfast reference.

I was lucky indeed to have his enthusiastic support and help, for it next occurred to me that it might be possible to date a number of surviving buildings by the use of the building covenants in the Donegall estate leases, of which first John Bristow and then my own firm had handed over several hundred to the Public Record Office. Ken Darwin, stretching to its utmost limit the agreement governing the deposit, allowed me to take home in batches the original counterpart leases. I took home also, this time from our own office, the bound map-book of Belfast that had been prepared and printed for the purpose of the sales through the Incumbered Estates Court. There, working at home on the dining-room table in Coolavin, the big house on whose site Queen's University subsequently built its disagreeable halls of residence, I laboriously transferred in pencil the basic information contained in the leases onto the printed maps.

The Leases

The third earl of Donegall, proprietor of Belfast, was killed at the siege of Monjuich, outside Barcelona, in 1706. In 1708, Belfast Castle burned to the ground. From 1706 until 1757 the town belonged to the fourth earl, who was incapable of managing his affairs, which were looked after by trustees. Throughout this period, short leases and tenancies, providing no incentive to build or repair, were the rule. The fifth earl, who succeeded his uncle in 1757, and his advisers, adopted a much more constructive policy. A careful survey of the whole town was carried out and old leases were either surrendered as they fell in, or exchanged for new ones. So on 20 July 1767 new leases, of which the counterparts for no fewer than 258 survive in the Public Record Office, were granted for almost every holding in the town. It is probable that a good many more survive in the bundles of title-deeds relating to other properties in inner Belfast.

It is very remarkable that so many of these indentures have survived. An indenture is defined by the *Oxford English Dictionary* as ‘a deed between two or more parties with mutual covenants, executed in two or more copies, all having their tops or edges correspondingly indented or serrated for identification and security’. This was originally intended to assist the illiterate to identify their own documents, a procedure that lasted into my own young days in practice. But almost all such documents were superseded by the conveyances executed by the land judges under the Incumbered Estates Act or by private sales so that their retention was, in almost every case, altogether pointless. But, as has already been remarked, lawyers have – and have always had – an extreme disinclination to destroy or discard documents that might yet be needed.

A fairly typical example of a Belfast building lease of 10 July 1767 is that granted to Elizabeth Byrtt, widow (Fig. 6).³ ‘Byrtt’ is a not uncommon corruption of ‘Brett’; I have sometimes wondered if this lady was not a relation of mine; but my family tree affords her no perch. The plot granted to her was not in her own possession, but rented out by her to Elizabeth Banks, widow, her under-tenant. It comprised a plot, with a frontage to Castle Street (at the head of High Street) of 25 feet, extending backwards ‘towards’ Rosemary Lane 135 feet. The plan endorsed on the appropriate space on the printed lease shows an L-shaped house (i.e. with a return) on the Castle Street frontage, a modest yard, a large cow-house occupying the entire width of the plot, a small out-house presumably a privy, and a long crooked back yard where it is to be supposed that the cow (or cows) took the air. The plot is bounded on the east by premises of Samuel Hyde, on the west by premises of Widow Cadell and to the rear by premises of Widow Dalzell (see map 21, plot nos 2, 3, 4). Widows appear to have been plentiful in Belfast at this period and over thirty are listed in the directory of leaseholders in part III. There seems to have been no access or right of way at the rear to Rosemary Street. Presumably the cow (or cows) got to the yard either by a gate at the side of the house – though none is shown on the plan – or through the house itself, an arrangement by no means inconceivable at that date. Six

³ PRONI, D509/191.

houses in Rosemary Street had been ‘lately built’ in 1767 immediately to the rear of Mrs Byrnt’s plot.

The lease was expressed to be for the lives of Arthur, earl of Donegall, his wife Ann, countess of Donegall and the honorable John Chichester, brother to the said earl and then the residue of a term of ninety-nine years from 1 May last past. The rent was to be £6 13s plus 2s duties payable on the usual gale-days of 1 May and 1 November in each year, plus a heriot of 5s on the death of the tenant and leet-silver at courts-leet. Mrs Byrnt further covenanted within fourteen years to:

“take down the messuage or tenement now standing”,

and on the frontage to Castle Street:

“in the stead or place thereof, in a good and workman-like manner, and with good sound materials, erect, build and finish . . . on the whole front of the said ground . . . one good and substantial messuage or tenement of brick and lime, or stone and lime, and shall cause the front of the said messuage to be sashed, and the roof thereof to be well slated, and all the side-walls of the said messuage to be built twenty-eight feet high above the surface of the ground, and fourteen inches thick at least, and that all the girders, joists and roofing, and the other timbers which shall be used or put therein, shall be of good oak or fir, and of proper scantlings respectively.”

She was also required to pave the street outside with stones or pebbles out to the middle of the street, and to keep it “well swept and cleansed”. Moreover, should she, or any of her tenants “lay or put upon any part of the streets of the said town of Belfast, any ashes, dung, filth or dirt” and allow it to remain there for twenty-four hours, she was to be liable to a fine of a shilling a day. There were also the usual covenants for repair, due payment, and for the (compulsory) use of the manorial corn mill.

By way of parenthesis, it is relevant to note that this was part of the site of the old Ulster Club, a noble building erected to the designs of Charles Lanyon in 1863. The club fell on hard times; it was ingeniously suggested by the late Brian Rankin that the back quarters should be demolished to provide offices for the Law Society, the chartered accountants, and the Royal Society of Ulster Architects on three floors, the fine formal rooms on the frontage to be shared by the three bodies. Unhappily, this scheme proved to be impracticable because of the lack of any rear access. So it was sold to a developer, who deliberately allowed it to fall into disrepair. After it had been vacated by the clubmen, it suffered slight bomb damage, but enough to allow the weather and the pigeons to penetrate the roof and so permit an outbreak of irremediable dry rot. It was until lately the site of the architecturally ignominious central Post Office.

Twenty years later, a lease in a somewhat different (but still printed) format was granted on 6 June 1790 by Lord Donegall to James Cooper, of Belfast, publican, for a fixed term of ninety-nine years, over an irregularly-shaped plot of land at the top of Donegall Street, opposite the poor house, with frontages to Carrick Hill (otherwise called ‘the Carrickfergus Road’) and

Patrick Street, at a rent of £14 per annum.⁴ Mr Cooper covenanted to build, within two years, three or more good and substantial houses of three storeys and that “the cornices doors and windows thereof shall range in a line with the houses nearest the church of Belfast standing in the same street”. Needless to say, this was the parish church of St Ann, not the much later St Patrick’s Roman Catholic Church almost next door. I had at the time of writing my book assumed that the building covenant had been duly complied with since there was a stringent proviso for forfeiture in the case of non-compliance. But it seems that this was not so, for when in 1989 the Hearth Housing Association was restoring the houses next to the parochial house, assumed to have been of 1809, it was discovered that the parochial house was in fact older than its neighbours, which presumably, therefore, were not built until the 1820s or 1830s.

On 20 April 1786, Lord Donegall granted a lease of a substantial site on the west side of his fashionable new street, Linen Hall Street, now Donegall Place, to his faithful architect and developer, Roger Mulholland, with a covenant to build upon it three very grand houses. That lease has not survived in the Public Record Office but of the three counterpart sub-leases, that to the Reverend Edward Patterson, clerk, curate in the parish of Belfast, has survived. His neighbours were, on the one side, John Smith, a prosperous merchant and, on the other side, the redoubtable Dr James McDonnell of the Glens of Antrim.

Mr Patterson’s sub-lease is dated 2 May 1791.⁵ The lease was for the fixed term of twenty-one years from 1 November 1791 at the formidable rent of £34 2s 6d per annum. In fact, he seems to have assigned it to James Luke in 1800, after only nine years. The printed document was a crisp, concise and straightforward one. It contained no unusual covenants. But it did contain a hand-written proviso, after the usual landlord’s covenant for quiet possession, as follows:

“and also that he the said Roger Mulholland his executors administrators or assigns shall and will allow the said Edward Patterson his executors administrators or assigns a sum of £11 16s in payment of the last half years Rent that shall be due and payable at the expiration of this demise, being the one half of the expence of all locks grates chimney pieces and shelves in said messuage and tenement”.

This house, now encased in mid nineteenth-century stucco, appears to be that still standing at no. 25 Donegall Place, though certainly now almost unrecognisable.

On 5 June 1800, Lord Donegall (and his trustees) granted a lease to one John Taylor of Belfast, cooper, of building ground off ‘the Back Plantation’ (Corporation Street) for three lives (as usual, all members of the Chichester family) and ninety-nine years at a rent of £12 15s covenanting to build two or more two-storey houses, this time subject to the further covenant that there should not be permitted on the premises “the trade of a brewer butcher slaughterer tobacco pipe binner tallow chandler melter of tallow brazier glass

⁴ PRONI, D509/758.

⁵ PRONI, D509/805.

maker founderer or any noisome or offensive trade whatsoever without the consent of the said marquis” (see map 2, plot no. 124).⁶ It is surprising that the notoriously smelly trade of tanner was not included, for Belfast was full of tanneries.

I hope that these examples will serve to give the reader some flavour of the leases whose particulars have been summarised on the maps and that their arcane phraseology will not be too difficult for a lay reader, familiar and even common-place as they must seem to every lawyer who has ever had to read and parse an eighteenth-century title.

The Incumbered Estates Court

In consequence of his debts, Lord Donegall could not continue to live on the revenue from his rents without meeting the claims of his creditors and, under the new legislation, he was obliged (in the first instance, by reason of the petition for sale lodged by a Mr Turner) to offer his Belfast estate for sale through the Incumbered Estates Court. The Incumbered Estates Act of 1849 was one of the most brilliantly successful pieces of legislation ever passed. It is remarkable that it has attracted so little attention from legal historians. The two most useful accounts of it, very different indeed from each other though they are, were written quite soon after the court, established in 1849, was wound up ten years later having accomplished its initial task. It was succeeded by an act of 1868, which established a successor body, the Landed Estates Court, with greatly extended powers. Incidentally, the name is often spelled Encumbered; but Incumbered is the correct, parliamentary, spelling.

The first account of its work is contained in the very dry, statistical appendix six⁷ to Richard MacNevin’s book, *The practice of the Landed Estates Court in Ireland*, published in 1859, and contains a summary of the proceedings of its predecessor from 25 October 1849 until 25 October 1859. The second account, jovially facetious but none the less equally useful, by one Percy Fitzgerald, Esq., MRIA, made its first appearance in 1861 in Charles Dickens’s magazine *All the year round* and was subsequently published as a small hardback booklet in London in 1862.

Percy Fitzgerald (1834–1925) was an odd fish: writer, sculptor, painter, barrister and a friend of Dickens. Richard MacNevin was one of a large family of rather intellectual Dublin solicitors, all Jesuit-educated, some at Clongowes, he at Oscott. Other contributions, dealing with particular aspects of the act, are to be found in Dr Maguire’s case history of Lord Donegall in the *Economic History Review*, 1976;⁸ in Mary Cecelia Lyons’s book of 1993 on the illustrations to the rentals;⁹ and in Pdraig Lane’s account.¹⁰

⁶ PRONI, D509/1360.

⁷ R.C. MacNevin, *The practice of the Landed Estates Court in Ireland* (Dublin, 1859), pp.439–49.

⁸ W.A. Maguire, ‘Lord Donegall and the sale of Belfast’ pp.570–84.

⁹ M.C. Lyons, *Illustrated Incumbered Estates, Ireland, 1850–1905* (Whitegate, 1993).

¹⁰ P.G. Lane, ‘The Encumbered Estates Court, Ireland, 1848–9’, in *Economic and Social Review*, iii (1972), pp.413–53.

Irish land law has always been complex and difficult. The crown grants of the seventeenth century (comparable to those providing the basis for title on the eastern seaboard of America) had been superimposed on a version of Anglo-Norman feudal law, in turn superimposed on the much older traditional native Irish *brehon* law. Land law was interpreted and administered by the courts of chancery. Their procedures, in Ireland as in England, were inscrutable and interminable. Dickens's case of *Jarndyce v. Jarndyce* was wholly plausible, no exaggeration at all. Lord Mansfield, in the late eighteenth century, allowed himself to become an Irish mortgagee; his case in chancery for the recovery of his money was not concluded until 1853, sixty years later. But Ireland had none the less one great advantage over England. In the reign of Queen Anne, in 1707, parliament had created a Registry of Deeds. No dealing with land, intended to be effective for more than twenty-one years, was thereafter to be enforceable unless it had first been registered by means of a memorial summarising its terms. The Deeds Registry of Ireland still subsists, in Henrietta Street, Dublin, and there one may (not without both difficulty and dust) consult the memorials of every such transaction in land in Ireland for the past three centuries: as I have myself done, though few others, save law searchers and intrepid local historians, ever do so.

Dissolute landowners such as the earl of Donegall borrowed far beyond their means, 'incumbering' their lands with mortgages, often in favour of English lenders, by way of security for repayment. But many, quite innocent, smaller and more provident landowners had, in times of some degree of prosperity, 'incumbered' their lands by creating charges to secure widows' jointures, or annuities for younger children, or provision for spinster aunts, or, indeed, for the support of the charities they had founded. All this elaborate structure crashed when the Famine came in 1845. Nobody could pay rent to his landlord if he had no income; nobody could pay interest to his lender, or income to his auntie, if he had no income from his rents. The result was, as it is today called, negative equity: disaster all round.

The remedy was in the hands of the Westminster parliament and, largely at the instance of the ex-Prime Minister, Sir Robert Peel, the remedy was applied. True, the first attempt, in 1848, was ineffectual, for it made the grave error of placing reform in the hands of the chancery courts. But the act of 1849 was more successful. It took the whole business out of the hands of the chancery courts and confided it to three full-time Incumbered Estates Court commissioners, with quite extraordinary powers and with no right of appeal from their decisions.

Under the new procedure any creditor who could show that the interest on his debtor's incumbrances exceeded one-half of the income of the estate might present a petition for sale to the commissioners. The procedure to be followed against 'the incumbered nobleman' is nicely described by Percy Fitzgerald.¹¹ First, an order for sale was made, to which the nobleman was invited to object if he will. When he failed to do so, the petitioner's solicitors inquired as to the affairs of the estate.

¹¹ Percy Fitzgerald, *The story of the Incumbered Estates Court* (London, 1862), pp.21–43.

“By-and-by all these labours of the eminent firm result most unexpectedly in a handsome folio volume, elegantly printed and copiously illustrated with lithographic plans, vividly-coloured drawings, sections and elevations, together with tabulated columns showing the tenancies, rents, and acreage – in short, such a topographical picture in one volume – of his estate as must have astonished the incumbered nobleman himself”.

And that is just what we have, for Belfast, although it is divided into map-book and rental.

There were, of course, maps or plans of each plot already endorsed on the leases and their counterparts, but new maps showing the whole estate were none the less required. A sum of 3d per statute acre was allowed by the court for surveying and lithography. Although many of the other maps appear to have been copied and enlarged from the most recent Ordnance Survey plans, those for Belfast were produced specially by the Dublin booksellers, Hodges and Smith, in 1850.¹² The actual work would presumably have been carried out by a professional surveyor hired by that company, since the most detailed town plans then existing were the 1830s Ordnance Survey maps at a scale of 1:2376, whereas the majority of the Hodges and Smith plans are at larger scales, ranging from 1:800 to 1:5040.

The impending sale by auction is then advertised, and after two or three months, duly takes place, one of the three commissioners usually acting as auctioneer, announcing to the astonishment of those accustomed to the ways of the chancery courts:

“The purchaser can have his conveyance executed, sealed, and delivered this very day! . . . and it will be satisfactory for him to know that a very small box indeed will hold the conveyance! . . . A small box! . . .”

The printed form of conveyance barely fills twenty lines, or half a page of duodecimo print. By-and-by, it was expanded into a single skin of parchment, which even included a map. Fourteen days of grace was then allowed for the purchaser to lodge his money in the Bank of Ireland. The money is then distributed by the commissioner among the creditors in order of their priorities, with any balance left over remitted to the incumbered nobleman.¹³

So what was the outcome of all this endeavour? To quote Percy Fitzgerald once again, if a little selectively,

“nearly two millions of acres, or about one-seventh of the available surface of the country, has been disposed of by public auction . . . nearly four-and-twenty millions sterling have been paid into the hands of the unflinching triumvirs, who nicely weighed and determined conflicting claims, representing a sum of two-and-twenty millions. Nearly 4,000 petitions from creditors have been presented, praying for a

¹² J.H. Andrews, *Plantation acres* (Belfast, 1985), pp.400–01.

¹³ Fitzgerald, *op. cit.*, p.41.

sale; 8,000 estates have been brought to the hammer; and some 4,000 titles have been perused by the triumvirs themselves . . . each personally waded through those dirty waves of vellum and faded yellow paper on which the true title to an estate usually drifts down . . . The old-fashioned chancery dilly [i.e. diligence] . . . rumbled on at a slow walk, and was ten years distributing a million sterling. The new legislative engine dashes by, express, and scatters four-and-twenty millions within the same space”.

By the end of its life, the court’s task had been completed. Fitzgerald records that, in the last months of its existence, it had no more than seventeen cases to deal with, that is, about four-and-a-half to each judge.

“There was nothing left for it to sell; there were no more patient mortgagees, exasperated by long suffering, to petition. Everybody was paid. Nobody was incumbered of land.”¹⁴

The same facts are more prosaically presented in the ninth annual report of the commissioners:¹⁵

1.	Number of petitions presented	4,413
2.	Number of absolute orders for sale	3,547
3.	Number of matters in which owners presented petitions	1,363
4.	Number of matters in which owners were bankrupt or insolvent	365
5.	Number of conveyances executed by the commissioners	8,364
6.	Number of estates sold by provincial auction	409
7.	Number of lots sold	11,024
8.	Number of boxes containing upwards of 300,000 documents and muniments of title, deposited in the Records Office	2,986
9.	Number of cases which had been pending in the court of chancery before being brought into the Incumbered Estates Court	1,298
10.	Number of Irish purchasers	8,258
11.	Number of English, Scotch, and foreign purchasers	324
12.	Amount of purchase money paid by English, Scotch, and foreign purchasers	£3,160,244
13.	Gross proceeds of sale to 31 August, 1858	£23,161,093

(See *note* overleaf)

¹⁴ *ibid.*, p.68.

¹⁵ ‘Summary of the proceedings of the Incumbered Estates Court, from the filing of the first petition, viz., 25 October 1849 to 31 August 1858, being the termination of the last sitting of the commission’ in MacNevin, *op. cit.*, pp.441–4.

*Note: the acreage of the land sold is estimated at about 2,800,000, and the rental at about £1,500,000.*¹⁶

The three principal objectives of the legislation had been to facilitate the speedy sale and transfer of Irish land, to simplify the means of doing so, and to introduce new capital into Irish property and agriculture. In the first two objects, the results were spectacularly successful; in the third, not so, for English and other investors proved unexpectedly reluctant to purchase land in Ireland. This was not an unmixed blessing. Many sceptics had forecast a kind of new settlement, with English and Scottish landowners overwhelming the indigenous peasantry. Perhaps surprisingly, it turned out that there was already a sufficient middle and professional class in Ireland to acquire, and pay cash for, all this land newly thrown on the market. The new proprietors were not always kinder landlords than their predecessors. A not undesirable rationalisation of the management of land took place, no doubt leading to some hardship, but at least the land remained in Irish hands.

How did this happy result come about? Principally, I think, through the diligence and hard labour of the commissioners themselves. One commissioner, with his examiner and supporting staff, handled each case from beginning to end – the redoubtable Mountifort Longfield himself looked after Lord Donegall. Two commissioners stand out in particular. Mountifort Longfield, believed by some to have been the brains behind the whole scheme, was the son of a Co. Cork clergyman. He was called to the Irish bar at the age of 26, but did not practise. He became first professor of political economy, then regius professor of feudal and English law in Dublin. The *Dictionary of National Biography* says that “he was esteemed an especially learned real-property lawyer . . . an active liberal, and assisted to draft the Irish measures of the first and second Gladstone administrations”. He died, aged 82, in 1884.¹⁷ His colleague, Charles James Hargreave, was equally academic: an Englishman and a mathematician, as well as a barrister and professor of jurisprudence, who moved to Dublin in 1849 on his appointment as a commissioner. Of him, the *Dictionary of National Biography* says “his mathematical essays were numerous . . . want of rest brought on an exhaustion of the brain, from which he died at Bray, near Dublin, 23 April 1866” – at the early age of 46.¹⁸

The Irish practitioner of conveyancing, at any rate in my day, greeted the signature of either of these gentlemen with a welcoming recognition, for a deed signed by either of them was sure to constitute that paragon among muniments, ‘a good root of title’, parliament having conferred upon Incumbered Estates Court conveyances, as on their successors Landed Estates Court conveyances, the very rare privilege of indefeasibility, meaning that no challenge whatever could be mounted to their validity once safely sealed, signed, delivered, and registered in the Registry of Deeds.

As Lord Justice Christian put it in a case of 1869:

¹⁶ These statistics have been slightly summarised and abbreviated; take into account that all these figures are at mid-Victorian money values.

¹⁷ *Dictionary of national biography*, xxxiv, pp.119–20.

¹⁸ *ibid.* (1890), xxxiv, pp.379–80.

“by a sort of conveyancing magnetism [an Incumbered Estates Court conveyance] . . . would draw out, not merely from the owner whose estate was under sale, or from whatever persons might intervene as parties in the proceedings, but from the absent, the helpless, the infant, the married woman, the mentally imbecile, nay, even the unborn, every particle of estate and interest, legal or equitable, present and future, known or unknown, latent or patent, in the land expressed to be conveyed, and would concentrate the whole in the purchaser, freed from everything that the conveyance did not save”.¹⁹

This uncommon privilege was upheld by the court of appeal on several occasions. There were a few mistakes, mostly of mapping and a few equitable injustices. But they were not many, even those incumbered noblemen who had at the outset complained so passionately about ‘robbery and confiscation’ in the end became quite contentedly reconciled to the outcome of the legislation.

There were one or two other commissioners, of whom the only one whose name is remembered was William Carey Dobbs. Each commissioner received a salary of £2,500 per annum in 1858. He was assisted by an examiner, at £600; a chief clerk, at £300; a second clerk, at £200; and a junior clerk, at £150 per annum. They were supported by 7 registrars, 4 accountants, 6 clerks of records and keepers of deeds, a taxing officer and his assistant, a stationery clerk, 3 criers, and 3 tipstiffs. The aggregate cost of these, with pensions to retired officials of the court, amounted to £18,500.²⁰ And that was the total annual salary and pensions bill for a staff responsible for sales worth £24,000,000!

As Percy Fitzgerald remarked,

“what may be done with five-and-twenty millions may surely be done with ten times that sum. There is a huge superficities in Great Britain, already handsomely burdened; there are mortgagees hungering and thirsting after their proper moneys, and labouring through the protracted formalities of the English court of chancery, to recover it. The cumbrous engines of that establishment are too slow and old-fashioned for the work of the age, even after all alterations and remodellings. They should be taken down, and new machinery put up with all convenient speed”.²¹

But it was not to happen, for the short and simple reason that to extend the Incumbered Estates code to England was wholly impracticable by reason of the complete absence, in that jurisdiction, of any equivalent of the Registration of Deeds, Ireland, Act, 1707. So there existed no public record of the incumbrances affecting any parcel of land.

¹⁹ *Re Tottenham’s Estate* (1869), IR 3Eq 528 at 547. I am indebted to Lord Carswell for this apposite citation.

²⁰ MacNevin, *Practice of the Landed Estates Court*, pp.440–41.

²¹ Fitzgerald, *Story of the Incumbered Estates Court*, pp.69–70.

In one other, perhaps rather surprising, location, where there did exist a system for the registration of incumbrances, the experiment was tried again: the West Indies. The abolition of slavery had had consequences there almost as disastrous as those of the Great Famine in Ireland. To quote from the report presented to parliament in 1884,

“from a variety of external causes, over which the planters had no control, the price of sugar, which had averaged £60 a ton in the early part of this century, had in 1850 fallen to £25. Any concomitant decrease in the expenses of cultivation and manufacture had been more than neutralised by the complete disorganisation of the labour supply following on the emancipation of the slaves, for the liberal compensation given to the planters on that occasion did not provide any labour in substitution . . . in the early years of the century, when sugar was at such a high price, planters had provided liberally for all dependent on them by jointures, annuities, legacies, and so forth, all chargeable on the estates. Under the new conditions, the planters in order to carry on the cultivation at all, were forced to borrow heavily and to burden with fresh incumbrances their already overburdened properties. The sugar planting industry was completely demoralised and estate affairs were hopelessly complicated”.²²

Unfortunately, the legislation in the West Indies enjoyed only very limited success. True, 290 estates, valued in total at a mere £445,950, passed through the court, and were “furnished afresh with unimpeachable titles” and “cleared of complicated incumbrances”. But this did not remedy the fundamental underlying lack of labour to replace the slaves. Since there was no available source of labour, so there were few purchasers. Accordingly, the Incumbered Estates legislation pioneered in Ireland proved in the end not to be a commodity susceptible to export elsewhere.

One final afterword on this subject. The legislation greatly simplified the affairs of the marquess of Donegall and his family and it gave very considerable satisfaction to his very numerous creditors, though, most unfortunately, the detailed records of the commissioners for sales and receipts appear not to have survived and it is not possible to give a reliable summary of the outcome. Working from the Donegall estate records alone, Dr Maguire notes that “it is difficult to make any accurate estimate of the amount raised by the sales; the landlord himself found it difficult. Lord Donegall’s reckoning in November 1853 was that about £11,000 of the former rental of £28,000 had been sold, which should have provided “upwards of £300,000”. But Verner [Lord Donegall’s agent] seems to have disagreed with these figures. By the end of the following year the annual income remaining was said to be nearer £16,000 than £20,000”.²³

Nevertheless, it can be said that, as Dr Maguire has shown, contrary to the earlier views of many local historians (amongst whom I must include myself), Belfast was not converted almost overnight from a town of tenants

²² *Report . . . on the West Indian Incumbered Estates Court* (London, 1884), p.4.

²³ Maguire, ‘Lord Donegall and the sale of Belfast’, p.581.

with short leases to a town of freeholders. So far as freedom to build was concerned, the landlord's control had virtually ceased by 1830 with the granting of most of the town and its neighbourhood in perpetuities. The building development which one would have expected to follow actually took place in the later 1820s' and the 1830s', as evidenced by the many late Georgian building leases recorded in this book. Fortunately for the Donegall family, its fortunes were considerably restored when in 1857 the sister of the late earl of Belfast married the son and heir of the wealthy, philanthropic, and above all respectable earl of Shaftesbury.

BOARD DIVERSITY: CAN SEX DISCRIMINATION LAW HELP?

Alice Belcher, Department of Law, University of Dundee

ABSTRACT

This article is about strategies that could be used to increase the number of women on the boards of large UK companies. There are two main ways in which the issue of women on boards has been tackled elsewhere: shareholder activism, for example in the US, and government imposed quotas, for example in Norway. The article investigates the possibility of using sex discrimination law to advance the position of women. It demonstrates the potential for a sex discrimination test case to be mounted in the UK. There are various problems that any such claim will face, both procedural and substantive, but it is not inconceivable that a case could be won. It is hoped that the credible threat of a sex discrimination claim can be added to pressures already being exerted to help bring about an increase in the number of women on UK boards.

Introduction

Statistics reveal very low proportions of women on UK boards of directors: The percentages of FTSE 100 boards with *no* women between 1999 and 2004 are shown in Table 1.

Table 1: FTSE 100 Boards with no women	
Year	Percentage of companies
1999	36
2000	42
2001	43
2002	39
2003	32
2004	31
Sources: 1999-2003; Singh and Vinnicombe, Summary of the Female FTSE 2003, Cranfield School of Management. 2004; The 2004 Female FTSE Index Both sources can be accessed from the dti's Women & Equality Unit web site.	

In 2003 just 5.3 per cent of directors across 568 UK listed companies were women. Across Europe's leading companies only 5.8 per cent of directors are women and more than half of these companies have no women on the

Board. Table 2 shows European statistics reported by the Ethical Investment Research Service (EIRS).¹

Norway	18.2
Sweden	13.6
Denmark	8.9
Ireland	7.3
Germany	6.4
France	6.2
Austria	6.0
UK	5.8
Finland	5.6
Netherlands	4.7
Switzerland	4.2
Greece	3.7
Belgium / Lux	2.6
Italy	2.6
Spain	2.5
Portugal	1.5
Source: EIRS statistics cited in <i>Business Life</i> , November 2003, 17.	

¹ Statistics are being collected across the world for instance; in 1992 just 4.3 per cent of the seats on the boards of North Carolina's largest companies were occupied by women. By 2000 the percentage of women on the boards of companies in S&P's indices was 9.8. Also in 2000, 89 per cent of the S&P 500 companies had at least one woman on the board: *PR Newswire*, 1 December 2000. This compares favourably with only 58 per cent of the FTSE 100 companies having at least one woman on the board in the same year. However, and again in the year 2000, 67 per cent of the state of Georgia's 228 publicly held companies had no women on their boards: *Atlanta Journal and Constitution*, 18 June, 2000. A survey Fortune 500 found that women held 13 per cent of board seats in April 2003: *Catalyst Census of Women Board Directors*, 2003, available at www.catalystwomen.org. Further US statistics are also available from Caralyst. So, the position for the US as a whole seems better than that in the UK or most of the rest of Europe, but within the US there are states where women still hold very few board appointments. In India the top 30 companies have 351 board positions and only seven of these are filled by women. To make matters worse, three of these positions are held by one woman, so there are only five women in board positions in these 30 companies: *Business Line*, 1 October, 2002.

Although most of the available statistics reveal low numbers of women on boards, the very fact that they have been collected and reported shows that board diversity generally and women on boards in particular are issues of interest and concern. In recent years board diversity has also received increasing attention in the academic literature where there have been treatments of the topic suggesting that, both theoretically and according to the empirical evidence, a diverse board can out perform a less diverse one.² In this article it is assumed that, in accordance with the majority of the literature, the goal of achieving a gender balance is a worthy one.

This article is about strategies that could be used to increase the number of women on the boards of large UK companies. There are two main ways in which the issue of women on boards has been tackled elsewhere: shareholder activism, for example in the US, and government imposed quotas, for example in Norway. These two strategies will be discussed briefly in the next two sections of the article. However, the main thrust of the article is to investigate the possibility of using sex discrimination law, in the form of a test case, to advance the position of women. This locates the article within a wider literature on women's participation in high level decision-making, for instance as members of parliament. In the field of politics, the Labour Party famously attempted to increase the number of women candidates by the use of women-only shortlists. This attempt was thwarted by the use of sex discrimination law. In *Jepson and Dyas-Elliot v The Labour Party*,³ the tribunal ruled that selection of candidates by a political party is covered by Part II of the Sex Discrimination Act 1975 headed up the "employment field". The use of women-only shortlists therefore amounted to unlawful positive discrimination. The response to this ruling was the Sex Discrimination (Election Candidates) Act 2002 which excludes candidate selection from the usual UK sex discrimination rules.⁴ Whilst this article concentrates on a specific form of sex discrimination test case, envisaged in the particular context of women as potential board members, it is also a contribution to the more general debate about how to increase the number of women in high level, overtly powerful jobs. In *Jepson* the UK's sex discrimination law was used by men resisting women-only shortlists. The response to *Jepson* was to treat election candidates as special cases. This response did nothing for the cause of women in areas other than politics. In this article the idea of women continuing to use the UK's sex discrimination law as a weapon is explored. The article explores the current mechanisms for the appointment of directors, then in its main analysis section, it investigates the possibility of a woman making a claim under the Sex Discrimination Act 1975 if she is not appointed to a UK board. Despite the problems that are raised by this analysis, the use of a sex discrimination test

² See for example Bilimoria and Wheeler "Women corporate directors: Current research and future directions" in Burke and Mattis (eds.), *Women in Management: Current Research Issues*, vol. II, (2000) ch.10, 138-163; and Singh, Vinnicombe and Johnson "Women Directors on Top UK Boards" (2001) 9 (3) *Corporate Governance: An International Review* 206-216.

³ [1996] IRLR 116.

⁴ This followed the publication of a report prepared by the UCL Constitution Unit; Russell *Women's Representation in UK Politics: What can be done within the Law?* June 2000.

case remains a possibility. The article concludes with some observations on the use of the language of 'board diversity' and of 'discrimination', including the credible threat of a sex discrimination claim, as a way of exerting pressure on UK companies.

Shareholder Activism

Shareholder proposals, on any topic, reaching the agendas of company annual general meetings (AGMs) are much more common in the US than in the UK. Some of these are championed by individual shareholders, but many are put forward by institutional shareholders. Some resolutions are sponsored by coalitions of institutional shareholders such as the resolution at EMC Corp's 2002 meeting calling for the company to take greater steps to create a more diverse and independent board which was sponsored by a coalition of 20 institutional shareholders. In 1994, Calvert Group of Bethesda, Md., an investment house, launched its first board diversity resolution and in 2003 it launched a record 9 such resolutions.⁵ Not all board diversity proposals are directed at gender inequality. In the US there is an equal concern about the small numbers of ethnic minorities on boards. In 2000 a proposal was put to the AGM of SunTrust, a Georgian-based Fortune 500 company calling for the board nominating committee to locate qualified women and people of colour. This proposal generated 13 per cent of the votes.⁶ The *Social Issues Reporter* publishes annual tables of the top vote-getting social policy resolutions. In 2001 these mostly focused on global labour issues. However, a resolution put at Unocal's meeting called for the board to report on and commit to board diversity and a resolution put to MBNA's meeting called for the board to report on steps to break the glass ceiling.⁷ In 2002 there were more proposals concerning board diversity with 9 of the 2002 resolutions asking companies to make greater efforts to diversify their boards or commit to board diversity. Most shareholder resolutions are put forward to draw attention to an issue with no real prospect of success. However, the EMC Corp coalition-sponsored resolution referred to above was passed with 56 per cent of the votes cast. Following the passing of the resolution in May 2002, the company elected a woman to its board in July 2002.⁸ Institutional shareholders in the UK have to date been more likely to exert influence through their investment policies. Investors holding large blocks of shares are also able to express their views directly to the company's management in private meetings.

Government Intervention

The Norwegian Government has tackled the issue of gender equality on company boards by imposing a quota. Norway has proposed a rule that 40 per cent of board members are to be women by 2005. A bill was presented to the Storting on 13 June 2003 amending company law relating to the composition of boards in all state-owned enterprises and corresponding

⁵ *Dallas Morning News*, 18 May, 2003.

⁶ *Atlanta Journal and Constitution*, 18 June, 2000.

⁷ *Social Issues Reporter*, July, 2001.

⁸ *Providence Journal-Bulletin (Rhode Island)* 16 July, 2002.

company law relating to public limited companies. The rules are expressed in terms of the minimum number of each sex:

“If the board has two or three members, both sexes must be represented

If the board has four or five members, sex must have at least two representatives

If the board has six to eight members, each sex must have at least three representatives

If the board has nine members, each sex must have at least four representatives, and if the board has more than nine members, each sex must make up at least 40 per cent of the representatives.”⁹

The Norwegian Government has been working to improve gender representation on company boards since the spring of 2002. The 40 per cent target was set first for state-owned enterprises and by the time the bill was introduced the percentage of women on the boards of companies owned wholly by the State was 45.7 per cent.¹⁰ The rules applying to state-owned enterprises came into force on 1 January 2004 a date when compliance had already been achieved. The rules applying to public limited companies will not come into force if gender representation is achieved voluntarily. This is to be decided on the basis of statistics provided by the Register of Business Enterprises incorporating changes up to 1 July 2005. If these statistics, which should have been available in August 2005, reveal women to have 40 per cent of the board seats, the detailed provisions will not be brought into force.¹¹ The Norwegian provisions have therefore been described as a target for women’s representation of 40 per cent by 2005. If the target is missed and the rules come into force, they will be monitored and enforced by the Register of Business Enterprises. The Register of Business Enterprises will refuse to register a company board if its composition does not meet statutory requirements. A company which does not have a board that fulfils the statutory requirements may be dissolved by order of the court.¹² The Norwegian provisions make no rules for private limited companies ‘. . . because most of these companies in Norway are small family enterprises where the owners are themselves members of the board’.¹³ The Equal Treatment Directive¹⁴ is applicable to Norway due to its Treaty commitments as a member of the European Economic Area. However, Norway has shown itself to be willing to legislate at the margins in respect of gender issues. It

⁹ Norwegian Ministry of Children and Family Affairs, in English: “Balanced gender representation on company boards – information about the proposed changes”. Accessed at <http://odin.dep.no/bdf/english/topics/gendereq/004051-990347/dok-bn.html> on 3/10/05.

¹⁰ Ministry Press release dated 13 June 2003, <http://odin.dep.no/bdf/english/topics/gendereq/004051-990347/dok-bn.html>.

¹¹ At 3 October 2005 statistics were not available.

¹² Norwegian Ministry of Children and Family Affairs, in English: “Balanced gender representation on company boards – information about the proposed changes” <http://odin.dep.no/bdf/english/topics/gendereq/004051-990347/dok-bn.html>.

¹³ *ibid.*

¹⁴ Council Directive EEC 76 / 207.

currently has an action pending in the EFTA Court in respect of its policy of reserving a number of academic posts exclusively for women.¹⁵

The use of targets or quotas is problematic for the UK because of the way in which the Equal Treatment Directive¹⁶ is implemented in the UK in the Sex Discrimination Act 1975 as amended. The 1975 Act effectively outlaws positive discrimination.¹⁷ However, targets for increasing the proportion of women in sectors where they are underrepresented have been held to be compatible with EC law when introduced elsewhere in the EC.¹⁸ Legally binding target-setting as a method of moving towards gender equality on UK boards is not out of the question as a matter of law.¹⁹ However, it is probably out of the question in the UK as a matter of politics. If it is unlikely that the UK will consider changes in its sex discrimination law to achieve a gender balance via a form of positive discrimination, this article's investigation of the potential for the UK's current sex discrimination law to be used to achieve the same objective is important.

The Appointment of UK Directors

A survey for Enterprise Network reported in 2000 that 74 per cent of non-executive directors (NEDs) were selected through personal contacts. In 2003 Higgs found only four per cent had a formal interview and only one per cent had answered an advertisement.²⁰ There are several consultancies specialising in finding NEDs and fees typically range from £5,000 to the equivalent of one year's director's fees. Higgs consulted and then reported on the role and effectiveness of NEDs. Responses to Higgs' consultation paper included the opinion that: 'Recruitment is easy as all the head hunters have long lists of available Non-Executive Director candidates.'²¹ However it was also suggested that: "Most Nomination Committees rely on head

¹⁵ Action brought on 22 April 2002 by the EFTA Surveillance Authority against the Kingdom of Norway (Case E-1/02). Brought on the basis that "by maintaining in force a rule which reserves a number of academic posts exclusively for women, [the Kingdom of Norway] failed to fulfil its obligations under Articles 7 and 70 of the Agreement on the European Economic Area and Articles 2(1), 2(4) and 3(1) of the Act referred to in point 18 of Annex XVIII to the Agreement (Directive 76/207/EEC of 9 February 1976 on the implementation of the principle of equal treatment for men and women as regards access to employment, vocational training and promotion, and working conditions)."

¹⁶ *ibid.*

¹⁷ Except in relation to election candidates as noted earlier. Positive discrimination is, of course, used as a way of promoting equality in other jurisdictions, e.g. the US.

¹⁸ *Re Badeck and others* [2000] All ER (EC) 289.

¹⁹ It should be noted that the UK has recently obtained a derogation from the Equal Treatment Directive, regarding religious discrimination, in the form of a particular provision of Article 15 of Council Directive 2000/78/EC in respect of the recruitment policies put in place for the new police service of Northern Ireland and the recruitment of teachers to schools in Northern Ireland.

²⁰ Higgs, *Review of the Role and Effectiveness of Non-executive Directors* (January, 2003) The Stationary Office London, para. 10.5, available at www.dti.gov.uk/cld/non_exec_review.

²¹ Responses to Higgs' consultation paper can be found on the dti website at www.dti.gov.uk/cld/non_exec_review/higgsresonse_page1.htm.

hunters to search for non-executives. My experience is that these head hunters have limited vision as to suitable candidates.”²²

The Higgs Report itself says nomination committees should insist that their consultants look beyond the ‘usual suspects’.²³ It also resulted in some changes to the Code on Corporate Governance including a new Principle that: “There should be a formal, rigorous and transparent procedure for the appointment of new directors to the board.”²⁴ However, there was no recommendation that board appointments should be advertised. Instead the new version of the Code demands that “external advice or open advertising”²⁵ be used. The Higgs Report has been followed by the Tyson Report²⁶ which looks at ways of engaging with “broader pools of talent”, including more women, when appointing directors.

A Possible Sex Discrimination Claim

This analysis imagines a well-qualified woman who is not appointed to a board of directors. She brings a claim of indirect sex discrimination. This work asks whether such a claim could succeed. The analysis is conducted at two levels. Firstly, there is the question of whether an employment tribunal has jurisdiction to hear such a claim at all, dealt with in the next subsection. Secondly, there are the substantive issues that could arise if such a claim were heard, dealt with in the following subsection. Once the test-case has got through the procedural hurdles the substantive issues appear to present fewer problems. However, success on the substantive issues will depend to a large extent on choosing a test case where the facts map easily onto the statutory wording of the ingredients of the claim.

Procedural Issues

If a company law text book is consulted about the appointment of directors it appears that they are voted into office by the shareholders at the annual general meeting (AGM). However, in practice most directors are co-opted by the board between AGMs and confirmed in their office by the shareholders’ vote. There are therefore three potential decisions that could be challenged under the Sex Discrimination Act 1975. It may be that the recommendation of the nomination committee could be challenged in the same way as the decisions of local Labour Party selection committees. There are a set of provisions in section 13 of the 1975 Act that catch the discriminatory acts of bodies or authorities that act as a filter through which applicants must pass. If a woman will only be considered by the board and proposed to the shareholders as a director on the recommendation of the nomination committee, this could be the place for a woman to launch her action.

²² *ibid.*

²³ Higgs (2003) *op. cit.*, para. 10.19.

²⁴ Code provision A.4.

²⁵ Code provision A.4.6.

²⁶ Tyson *The Tyson Report on the Recruitment and Development of Non-Executive Directors* (June, 2003), London Business School, available at www.dti.gov.uk/cld/non_exec_review.

In *Jepson and Dyas-Elliot v The Labour Party*,²⁷ the tribunal ruled that selection of candidates by a political party is covered by Part II of the Act headed up the ‘employment field’. However, a recent Scottish case shows how this sort of claim can be thwarted. In *Secretary of State for Scotland and the Advocate General for Scotland v Mann and McCourt*²⁸ the actions of a returning officer in the first election for the Scottish Parliament were challenged as discriminatory. The Employment Appeal Tribunal (EAT) held that a returning officer is an authority or body under section 13 and that membership of the Scottish Parliament constitutes an occupation. However, the judgment states: “acts and omissions of the returning officer are fenced by a criminal provision . . . the Westminster Parliament intended . . . to exclude him from attack within the civil justice system.” Therefore, the EAT held that the employment tribunal system had no jurisdiction to hear the case. This case opens the way for arguments about alternative methods of challenge to be presented as a way of escaping the clutches of discrimination law. If company law is the right way to challenge the choice of a company director, perhaps the employment tribunal could be convinced that, as in *Mann*, it should deny itself jurisdiction to hear the case.

Another reason why a claim of this sort may not succeed is the definition of bodies to which it applies. Section 13 of the Sex Discrimination Act 1975 applies only to: “. . . an authority or body which can confer an authorisation or qualification which is needed for, or facilitates, engagement in a particular profession or trade”. The nomination committee is always constituted as a sub-committee of the main board of directors. There is therefore an argument that it has no decision-making powers and cannot be viewed as “an authority or body”. Decisions of the board of directors may more easily be seen as decisions of “an authority or body”, however, it may be difficult to separate the decisions of the board from decisions of the company itself. The board is not a filter through which an applicant must pass before presenting themselves to their prospective employer, as it is the board that makes the employment decision when a new director is co-opted.

It may be, however, that the board’s decision to co-opt could be challenged as a decision of the company as ‘employer’. Under section 6 (1) of the Sex Discrimination Act 1975 it is unlawful to discriminate against a woman in relation to employment:

“(a) in the arrangements he makes for the purpose of determining who should be offered that employment, or . . .

(c) by refusing or deliberately omitting to offer her that employment.”

Employment is defined in section 82 (1) of the Act as “employment under a contract of service . . . or a contract personally to execute any work or labour”. This is the sort of claim envisaged in the substantive analysis conducted in the next subsection. The main difficulty with this claim is that the office of director does not automatically involve a contract of service, *i.e.* an employment contract. A woman has already brought a successful claim of sex discrimination on the grounds that she was not offered the same

²⁷ [1996] IRLR 116.

²⁸ EAT/56/00.

opportunities to rise to the position of director as a male colleague; *Love v Federal Mogul Aftermarket UK Ltd.*²⁹ However, this claim was based on the denial of an “opportunity for promotion” and the tribunal did not have to consider whether the office of director as such was covered by the Act. The tribunal only needed to establish that Love was in employment as defined and, presumably, that the phrase “access to opportunities for promotion” in section 6(2)(a) of the Sex Discrimination Act 1975 covered opportunities to rise to the position of director.

In a test case on the failure to appoint a woman as director, the tribunal would have jurisdiction to hear the claim if a director is employed under a contract of service *or* a contract personally to execute any work or labour. Executive directors are generally viewed as having employment contracts and as such the actions of their employers (the company) would appear to be actionable under the Sex Discrimination Act 1975 within the category “employment under a contract of service”. As a fall-back position, if the tribunal did not recognise an executive director’s position as being “employment under a contract of service”, failure to appoint to the position of executive director would also be actionable if executives have contracts “personally to execute any work or labour”. Non-executive directors do not have employment contracts although they do receive remuneration from the company. Failure to appoint a woman as a non-executive director would only be actionable if non-executives have contracts “personally to execute any work or labour”.

There are therefore two issues to investigate: Do executive directors have employment under a contract of service? Do directors (both executive and non-executive) have contracts personally to execute any work or labour? The question of whether directors can claim employment status for the purpose of benefiting from employment protection legislation has been addressed in several cases: *Buchan v Secretary of State for Employment*;³⁰ *Fleming v Secretary of State for Trade and Industry*;³¹ *Secretary of State for Trade and Industry v Bottrill*³² and *Connolly v Sellers Arenascene Ltd.*³³ These cases show that in determining the issue of employment status, tribunals should apply the established common law tests for distinguishing employees from independent contractors; tests such as the control test, the economic reality test and mutuality of obligation.³⁴ The problem giving rise to all these cases is the position of a controlling shareholder who is also an executive director. These cases show that whilst it cannot be taken for granted that an executive director will be found to have a contract of employment, the sort of executive director’s position to which a woman bringing a sex discrimination test case would be aspiring is unlikely to fall outside the employment status. The sort of position that this article has in mind is one where there would clearly be a mutuality of obligations in the form of work to be done for a salary and the

²⁹ (unreported) 1999.

³⁰ [1997] IRLR 80.

³¹ [1997] IRLR 682.

³² [1998] ICR 564.

³³ [2001] ICR 760.

³⁴ See Wardman ‘Directors and Employee Status: An examination of relevant Company Law and Employment Law Principles’ *Company Lawyer*, (2003), 24, (5), 139-143.

controlling shareholder issue would not arise. There are, however, other reasons why it might be better for a test case to be brought in relation to a non-executive directorship. In particular, it may be easier to argue that the claimant is qualified for the post if a non-executive vacancy is used as the non-executives are supposed to bring independence and, to some extent, an outsider's views to the board. It would correspondingly be harder for the company to establish that industry experience is a genuine occupational requirement in the case of a non-executive director.³⁵

If the test case is to focus on a non-executive directorship, then the relevant issue is whether non-executive directors are covered by section 82 (1) of the Act because they have contracts "personally to execute any work or labour". It may seem obvious that this is so. The company expects non-executive directors to attend board meetings in person, no doubt this can be viewed as work, and pays them for their efforts. There would appear to be no problem in demonstrating mutuality of obligations. However the issue may not be whether non-executives "personally execute any work or labour" but whether they do so under a "contract". The issue is whether the relationship between the non-executive director and the company is contractual or statutory. This issue has been explored in the context of the relationship between general practitioners and health authorities in a number of cases.

In the case of *North Essex Health Authority v David-John*³⁶ a General Practitioner claimed unfair dismissal and racial discrimination. The Employment Tribunal held that the GP was an employee under section 230 of the Employment Rights Act 1996 and was in 'employment' as defined in section 78 of the Race Relations Act 1976.³⁷ The Employment Tribunal therefore held that it had jurisdiction to go ahead and hear the substance of the unfair dismissal and race discrimination claims. The Health Authority appealed on the grounds that these findings were incorrect as a matter of law. The EAT held that the relationship between a GP and a Health Authority is purely statutory and not contractual: It is based on the National Health Service (General Medical Services) Regulations 1992.³⁸ The terms of service impose obligations on a GP in respect of patients, but these are statutory not contractual. Therefore the GP was not an employee under section 230 of the Employment Rights Act 1996. Also, because of the purely statutory nature of the obligations, the GP did not have a *contract* "personally to execute any work or labour" under section 78 of the Race Relations Act 1976.³⁹ The Employment Tribunal did not have jurisdiction to hear the unfair dismissal or race relations claims. What the tribunal should

³⁵ See next subsection for a discussion of this and other substantive issues.

³⁶ (unreported) 15 August 2003 EAT Appeal No EAT/0232/03/ILB.

³⁷ "Employment" is defined in s.78 of the RRA 1976 in identical terms to the definition found in s.82 of the Sex Discrimination Act 1975.

³⁸ These were the regulations applying at the time of Dr David-John's case. At the time of writing GPs are talking about their "new contract" – which no doubt has terms and obligations different from those laid down in the 1992 regulations, but is unlikely to be contractual in the sense required by the ERA.

³⁹ This is the equivalent of s. 82 of the Sex Discrimination Act 1975.

have done was to ask the question whether, in all the circumstances, the relationship was in fact contractual or purely statutory.⁴⁰

It was held in *Newtherapeutics Ltd v Katz*⁴¹ that the appointment of a person as a director of a company does not in itself form any contract between the person and the company. This case involved an application for leave to serve a writ on a director of an English registered company, the director being domiciled outside the jurisdiction in the US. Leave was sought under R.S.C., Ord., 11 to serve the writ and the court held that a claim against a director of a company based merely upon the fact of his appointment and in the absence of a specific contract of employment did not fall within the ambit of R.S.C., Ord., 11, r. 1(1)(d) (i) and (iii) which refers to a claim to “enforce . . . a contract, or to recover damages . . . in respect of the breach of contract.” Knox J. held that “I should be extending Ord., 11, r. 1(1)(d) (i) and (iii) to treat it as covering an appointment to office.” It is submitted, however that the facts of *Newtherapeutics Ltd v Katz* will be fairly easily distinguished from the facts that will pertain to the appointment of most non-executive directors of large UK companies and certainly to the target of any test sex discrimination case. The claim against such a company will not be based “merely upon the fact of the appointment”. Many of a director’s obligations are set out in companies legislation or are matters of common law and are therefore not negotiable, however, the Combined Code on Corporate Governance requires the nominating committee to “. . . evaluate the balance of skills, knowledge and experience on the board and, in the light of this evaluation, prepare a description of the role and capabilities required for a particular appointment”.⁴² Also, the Combined Code requires the terms and conditions of appointment of non-executive directors to be made available for inspection.⁴³ It is submitted that to the extent that there are any company- or post-specific understandings between the company and its non-executives these should be viewed as being of a contractual nature and as standing in addition to the general statutory obligations. Given the Combined Code provisions, there will always be some company- or post-specific terms and conditions involved in relation to listed companies. In this way *Newtherapeutics Ltd v Katz*, where the relevant claim was based merely on the appointment as a director, can be distinguished. In this way the GP – Health Authority position can also be distinguished. For a GP the relevant regulations constitute the whole of the legal relationship with the Health Authority (including the basis for calculating remuneration) whereas, for the non-executive director, statute and case law set out an irreducible minimum set of obligations which can be, and usually are, supplemented by contractual commitments between the director and the company. It is on the basis of a claim by a potential non-executive director; a post falling within the definition of employment by virtue of it being a contract personally to

⁴⁰ Other cases on the issue cited in *North Essex Health Authority v David-John* are: *Wadi v Cornwall and Isles of Scilly Family Practitioner Committee* [1985] ICR 492 and *Ealing Hammersmith and Hounslow FHS v Shukla* [1993] ICR 710.

⁴¹ [1991] Ch 226.

⁴² *Combined Code on Corporate Governance*, provision A.4.2 (Financial Reporting Council, July 2003).

⁴³ *ibid.*, Code Provision A.4.4.

execute work or labour, that the substantive issues are examined in the next subsection.

An alternative approach would be to challenge the decision that the shareholders make when they vote on the appointment at the AGM.⁴⁴ However, the shareholders are not usually offered a choice of candidates. They are usually faced with a resolution that *x* should be appointed to the board. The nature of their discriminatory act is therefore problematic. Also there are general rules of company law to be considered. For instance: ‘. . . the right to vote is attached to the share itself as an incident of property to be enjoyed and exercised for the owner’s personal advantage’; *Peters’ American Delicacy Co. v Heath*.⁴⁵ There are, however, some circumstances where company law demands that members vote ‘*bona fide* in the interests of the company as a whole.’ There is some authority suggesting that the appointment of a director may be such a circumstance but this is a moot point. The proposition can be traced to the speech of case of Jenkins LJ in the case of *Re H.R. Harmer Ltd.*⁴⁶ This case involved a family company where the father controlled a majority of the shares. Jenkins L.J. stated:

“There were in the course of the history of this unfortunate matter a remarkable number of appointments of directors and retirements of directors brought about in one way or another by the father. . . . It cannot be denied that the holder of the majority in voting power of the shares in a company may, broadly speaking, appoint any person he thinks fit as director, and the appointment cannot be challenged merely on the ground that he might have found some more suitable person than the person he selected, or that the person he selected was his friend; but I take it that the majority shareholder’s power of appointing directors must within broad limits be exercised for the benefit of the company as a whole and not to secure some ulterior advantage.”

He also stated:

“. . . the facts of this case are by no means usual, and I would not have it go forth that every time a majority shareholder appoints directors of his own choosing, he has done something wrong, or something which can be challenged by a dissatisfied minority, but if he goes on, having seen to it that the requisite majority is obtained, to state his motives for having Mr A rather than Mr B, and says: “Mr A will always vote in the way I tell him to,” then it seems to me it is impinging on dangerous ground.”

⁴⁴ The Companies Act 1985 says little about the appointment of directors and leaves this matter to the articles. If Table A articles are adopted, Art.78 empowers the members the appoint by an ordinary resolution (51 per cent). In practice many directors are co-opted by the board before having their appointment confirmed by a vote at the AGM.

⁴⁵ (1939) 61 CLR 457, at 504.

⁴⁶ [1959] 1 WLR 62.

The dictum has been expressed in short as “the majority shareholder’s power of appointing directors must within broad limits be exercised for the benefit of the company as a whole and not to secure some ulterior advantage”. However, when this phrase is quoted in context, it becomes clear that the case could be distinguished relatively easily. Also, the Company Law Steering Group (CLSG) considered this principle, but only as applying to the alteration of articles or class rights. The Group’s recommendation is that the area should be left for the courts to develop, but that the test itself should be codified as: “. . . whether the majority honestly believe that their vote is best calculated to promote the success of the company for the benefit of its members as a whole.”⁴⁷ The company law position must either be that the shareholders’ right to vote can be exercised for the owner’s personal advantage (*per Peters’ American Delicacy Co. v Heath*) or that the shareholders must vote in the way that they honestly believe is best calculated to promote the success of the company for the benefit of its members as a whole (per the CLSG formulation of the test). Neither of these company law positions requires a shareholder to pay attention to board diversity in a direct way. If board diversity is demonstrated to be associated with corporate success, and the duty to vote *bona fide* in the interest of the company applies in the circumstances, an argument in favour of shareholders having a duty to choose greater diversity over less diversity could be made. However, another part of the speech in *Re H.R. Harmer Ltd* should also be remembered: “the appointment cannot be challenged merely on the ground that he [the majority shareholder] might have found some more suitable person than the person he selected.”⁴⁸

Substantive Issues:

The current definition of indirect sex discrimination in the context of the employment field is set out in section 1(2)(b) of the Sex Discrimination Act 1975. A person discriminates against a woman if

“he applies to her a provision, criterion or practice which he applies or would apply equally to a man, but –
 which is such that it would be to the detriment of a considerably larger proportion of women than of men, and
 which he cannot show to be justifiable irrespective of the sex of the person to whom it is applied, and
 which is to her detriment.”

This definition was substituted by S.I. 2001/2660 reg 3 in place of the definition set out in the original 1975 statute; a definition that remains in place for sex discrimination outside the field of employment. This earlier definition states that a person discriminates against a woman if

⁴⁷ Final Report para 7.58.

⁴⁸ Who should bring such a claim and in what capacity is also a problem. The most obvious scenario is for a minority of shareholders to challenge the decision of a majority on the grounds that the majority have not voted *bona fide* in the interest of the company as a whole and they have a duty to do so. The potential woman director would therefore need to be championed by the minority shareholders.

“he applies to her a *requirement or condition* which he applies or would apply equally to a man but –

which is such that *the proportion of women who can comply with it is considerably smaller than the proportion of men who can comply with it*, and

which he cannot show to be justifiable irrespective of the sex of the person to whom it is applied, and

which is to her detriment *because she cannot comply with it.*”

The italics show how the old definition differs from the new.⁴⁹ The proposed test case for a woman not appointed as a non-executive director must be analysed using the new definition, but some of the relevant authorities are, of course, cases decided under the earlier wording. Where the change in the definition might affect the authority of earlier cases the impact of the new wording will be noted.

When searching for a new non executive director, companies often require someone with several years experience as an executive director. This practice will, given the current numbers of women executives, have a disparate impact. Also, it has been found that companies often use consultants to search for new directors. This mechanism has been used in the knowledge that the lists of potential directors held by such consultancies are disproportionately male. The condition that potential directors are recommended by consultancy *x*, where the list of names available is disproportionately male could therefore satisfy the indirect discrimination provision. Where a consultancy is not used, directors have often been appointed because they are personally known to other members of the board. This requirement would be a problematic one to be used as a test case for discrimination in the appointment of a director following the difficult case of *Jane Coker and Martha Osamor v The Lord Chancellor and the Lord Chancellor's Department*⁵⁰. *Coker* arose out the Lord Chancellor's appointment of a male special adviser, Garry Hart, a senior partner in Herbert Smith a firm of solicitors, who was personally known to him. When making the appointment, the Lord Chancellor did not advertise the post, and there was no open competition. He did not look outside his circle of acquaintances. The case was decided on the assumption that the Lord Chancellor had applied the “requirement or condition” (old wording) that the candidate be personally known to him. The point at issue on appeal was whether that condition had a disparate impact, as defined in the old wording. The court held that there was no disparate impact in the particular circumstances. It also opined that, in general, when appointment is through personal knowledge, disparate impact would not be demonstrated. The reasoning appears to have been as follows: Ask “what proportion of qualified women was excluded from consideration by the application of the condition that the person should be personally known to the Lord Chancellor?” – answer “100 per cent.” Ask “what proportion of qualified men were excluded from consideration by the application of that condition?” – answer “almost 100 per cent.” Can there be a significant difference between 100 per

⁴⁹ The old definition now appears as section 1(1)(b) of the SDA 1975.

⁵⁰ [2002] IRLR 80.

cent and almost 100 per cent? – answer “no”, therefore impact was not disparate. This case has been criticised and a similar case may produce a different result in the future, not least because of the opportunity offered by the changing of the disparate impact provision. Therefore, the decision in *Coker* may not remain a valid precedent.

For a test case perhaps the easiest provision, criterion or practice to attack might be the requirement for executive experience on another board. The Company could, however, argue under section 1(2)(b)(ii) of the Sex Discrimination Act 1975 that this is a justifiable requirement. The test for justification is a proportionality test developed by the EJC in the case of *Bilka-Kaufhaus GmbH v Karin Weber Von Hartz*.⁵¹ Under this test a practice is capable of justification by reference to a legitimate objective where the means chosen are appropriate and necessary to that end. Put another way; an exclusion must be based on objectively justified factors unrelated to any discrimination on the grounds of sex and this may include economic grounds.

The Irish case of *Patricia Conlan v University of Limerick and Minister for Enterprise, Trade and Employment (notice party)*⁵² has some parallels. This case concerned an advertised post of Professor of Law in the University of Limerick. The advertisement specified that the candidate would hold “a higher degree, preferably at doctorate level . . . have several years experience at a senior academic level and be a leading published researcher in a specialist field of law.” The plaintiff complained that the requirement to have several years’ experience at a senior academic level amounted to indirect sex discrimination. She argued that there would be very few women qualified for the position as there were very few women at a senior academic level. The University’s case was that the requirements were justified as they were essential in the circumstances. The Labour Court found the requirement to be objectively reasonable and essential. On appeal, although the High Court found that the Labour Court should have looked only at whether the requirement was essential, rather than “reasonable and essential”, it confirmed the finding in favour of the employer. This is an Irish case, but suggests that a company requiring new directors to have executive experience may be able to justify that requirement under the *Bilka-Kaufhaus* test. However, companies may find it more difficult to convince a tribunal that executive experience is essential in view of the suggestions of the Tyson report. Finding non-executive directors by looking only in the pool of executive directors of other UK plcs is the practice that is being attacked by Tyson’s recommendations that boards should look at broader pools of talent.

To summarise: Despite statistics that reveal massive under-representation of women on UK boards, a sex discrimination claim would not be straightforward. It has been argued that a claim could be made out most easily by attacking the decision of the board to co-opt as a decision of the employer under section 6 (1) (a) of the Sex Discrimination Act 1975. The main procedural challenge to this claim could be that the appointment of a director is a matter of company law and company law should be used to

⁵¹ (Case 170/84) [1986] ECR 1607, [1986] IRLR 317.

⁵² [1999] 2 ILRM 131.

challenge it.⁵³ There are also substantive issues, but it is submitted that a carefully prepared case arising out of circumstances auspicious to the claim could succeed.

Conclusions

The legal analysis offered in this article suggests that a test case based on UK sex discrimination law is likely to face obstacles. However, the language of discrimination is already being applied to the issue of board diversity. At the 2003 AGM of SouthTrust, Alabama's largest bank, a shareholder voiced concerns about the fact that the bank had no females in corporate management and only one on the board. It was suggested that the bank may be making itself vulnerable to a discrimination lawsuit similar to the one concerning discrimination against salaried black workers settled by Coca-Cola in 2000. The Coca-Cola suit was filed in April 1999 and claimed that the company discriminated against salaried black employees in pay, promotions and evaluations, settled in 2000 at a cost to Coca-Cola of \$192.5 million.

This article has discussed in detail the idea of mounting a sex discrimination test case as a way of changing the behaviour of companies selecting directors. It has also mentioned three other approaches; legislation requiring a quota to be attained; investor activism at AGMs and; encouraging boards to look at broader pools of talent. There are problems with the legislative solution in the UK context in terms of political acceptability and compatibility with EC law. Under any of the other approaches, which are not mutually exclusive, change may only be achieved if attitudes change, and attitudes can depend very much on the way language is used.⁵⁴ The first point to be made concerning the language of board diversity is a positive one. The use of the term 'board diversity' is itself an indication that the issue is a live one. The association of a board diversity problem with the threat of an expensive discrimination suit at the SouthTrust AGM is an explicit recognition of that lack of diversity may be discriminatory. The setting up of organisations to advance women in the business world also makes the issue more visible.⁵⁵

⁵³ In *Secretary of State for Scotland and the Advocate General for Scotland v Mann and McCourt* the EAT refused to hear a sex discrimination case because the mechanisms provided by the Scotland Act 1998 should have been used.

⁵⁴ The comments on the use of language which follow are intended to indicate positive and negative gendered aspects of the language currently used in relation to boards of directors. A detailed review of the literatures on, for instance, language and decision making or language and the construction of reality is not intended. One philosophical example claim is however offered: 'Language offers a mechanism for putting myself into the world, as Heidegger might phrase it, and for making the world part of me; and language very likely determines the way in which experience will be registered and later recalled.': Spence "Turning Happenings into Meanings: The central role of the self" in P.Y. Eisendraith and J.A. Hall (eds.) *The Book of the Self: Person, Pretext and Process* (1987), 134.

⁵⁵ For example, in the UK Cranfield School of Management has a Centre for Developing Women Business Leaders, in the US and Canada Catalyst is a non-profit research and advisory organisation – www.catalystwomen.org.

However, if language has the potential influence behaviour there is certainly room for improvement in the way language is used in and around UK boardrooms. If it is true that “codes of established narratives . . . define our capacities to tell our individual stories”,⁵⁶ the UK needs to escape from boardroom narratives of directors who are men and non-executives who are executives of other plcs. The strength of these narratives is illustrated by the fact that the Tyson Report, which to many outsiders seems to suggest the obvious, was needed at all.⁵⁷ The use of gender-free language is insisted upon in many universities, but a request for a gender-free language rule at a recent corporate governance conference was met with the question “what would that mean?”⁵⁸

To conclude, this article illustrates how the issue of the lack of women on boards has become more prominent over recent years. Statistics are now being gathered across the world and efforts are being made in various ways to address the problem. This article has demonstrated the potential for a sex discrimination test case to be mounted in the UK. There are various problems that any such claim will face, both procedural and substantive, but it is not inconceivable that a case could be won. It is hoped that the credible threat of a sex discrimination claim can be added to pressure from shareholders and from government commissioned reports by Higgs and Tyson to bring about an increase in the number of women on UK boards. However, it may be that change will only happen when “codes of established narratives” are changed, and with them attitudes.

⁵⁶ Benhabib *The Claims of Culture: equality and diversity in the global era*, (Princeton, Princeton University Press, 2002), 15.

⁵⁷ The Tyson Report makes recommendations for voluntary change by UK listed companies. It encourages them to look outside the normal pool of potential non-executive directors which has been limited to those already in very senior positions in other UK listed companies. It suggests looking in the public sector and internationally for potential non-executives.

⁵⁸ A suggestion made by the author at the Annual International Conference on Corporate Governance held at the Henley Management College, autumn 2002.

A CRITIQUE ON LEGAL ANALYSIS OF LOCAL GOVERNMENT AND THE CENTRAL-LOCAL RELATIONSHIP

*Barbara Mauthe, Lecturer in Law, Lancaster University**

Introduction

This paper will argue that analysis on local government and the central-local relationship is generally confined to context and that a consequence of this approach is that theory, within this area of scholarship, is viewed as being mutually incompatible. Specific examples of this incompatibility can be found in respect of the works of Loughlin, Vincent-Jones and Cooper. It will be argued that by juxtaposing the work of these three writers, it is possible to connect their analyses, beyond the context within which they write, but that such a connection cannot be achieved through the mechanism of theory. The connection can however be achieved through the use of concepts and the resulting analysis is a form of meta-explanation of a particular period of decision making in local government and the central-local relationship.

Local Government, The Central-Local Relationship And Public Law Analysis

Local government and its relationship with central government are topics which do not attract much attention within public law scholarship. This is possibly because these are areas which are perceived as not raising questions of a wider constitutional or theoretical nature since their focus is the internal institutional structure of the state. When contrasted with other areas of public law scholarship, such as regulation and its analysis of rules and rule making, particularly in the light of the changes to the formal structure of the state, or human rights and its focus upon the relationship between the citizen and the state, local government and the central-local relationship can appear irrelevant and even insular.

These perceptions are further enhanced by the manner in which the institution of local government is presented. Local government, for example, is generally represented as either a distinct institution¹ or in terms of a specific area of decision making, such as housing, education or welfare. Examination of these topics may consider matters such as privatisation, deregulation and human rights, but these issues are not the basis of the study. Generally, analysis is subject specific and is often concerned with particular contextual aspects and even technical detail. Theoretical concerns are

* I would like to thank Chris Willmore, Bristol University and Agata Fijalkowski, Lancaster University for their helpful comments on an earlier draft of this paper.

¹ For example, see J. Sharland, *A Practical Approach to Local Government Law* (London: Blackstone, 1997) or S. Bailey, *Cross on Principles of Local Government Law* (London: Sweet & Maxwell, 2nd ed., 1997).

generally excluded although when included, the theory is 'borrowed' and its use is also confined to the specific context or subject area concerned.²

Local Government, The Central-Local Relationship And Theory

Notwithstanding the technical and contextual emphasis within legal analysis on local government and the central-local relationship it is possible to identify a theoretical element. The works of Loughlin³, Vincent-Jones⁴ and Cooper⁵ are unusual in that, whilst varying attention is paid to detail and context, their analyses also address the question of the wider role of local government along with its relationship with central government. In other words, and importantly for the purposes of this paper, at the core of each analysis is theory.

Loughlin, for example, argues that in the 1980's the relationship between central government and local government underwent a process of 'juridification'. This juridification occurred by means of two distinct processes: legalization and restructuring. Vincent-Jones creates a model of central-local relationship entitled 'responsibilization'. This model stems from observations made in respect of the 'responsive benefits' which the policy of compulsory competitive tendering has brought to local authorities. Benefits, Vincent-Jones argues, which have resulted in the emergence of a form of 'contractual governance'. Cooper argues that there is real pressure on local authorities to disguise their decision making as being non-ideological. Accordingly many decisions are presented in terms of neutral factors, such as economics or managerialism. However, many decisions are

² For example, see D. Cowan *Housing Law and Policy* (London: Macmillan, 1999) where the theory of autopoiesis is discussed in the context of housing services.

³ M. Loughlin 'Law, Ideologies and the Political-Administrative System', [1989] *Journal of Law and Society*, Vol. 16, page 21; *Legality and Locality: The Role of Law in Central-Local Government* (Oxford: Clarendon Press, 1996); 'The Restructuring of Central-Local Relations' in J. Jowell and D. Oliver (eds.), *The Changing Constitution* (Oxford: Oxford University Press, 4th ed, 2000), 137.

⁴ P. Vincent-Jones 'The Limits of Near-contractual Governance: Local Authority Internal Trading Under CCT' [1994] *Journal of Law and Society* Vol. 21, page 214; 'Responsive Law and Governance in Public Service Provision: A Future for the Local Contracting State' [1998] *Modern Law Review* 362; 'The Regulation of Contractualisation of Quasi-Markets for Public Services' [1999] *Public Law*. 304; 'Competition and Contracting in the transition from CCT to Best Value: Towards a More Reflexive Regulation?' *Public Administration* Vol. 77, No 2 1999 (273-291); 'Contractual Governance: Institutional and Organizational Analysis' [2000] *Oxford Journal of Legal Studies* 317; 'Central-Local Relations under the Local Government Act 1999: A New Consensus' [2000] *Modern Law Review* 84; 'Values and Purpose in Government: Central-Local Relations in a Regulatory Perspective' [2002] *Journal of Law and Society* 27; 'Regulating Government by Contract: Towards a Public Law Framework' [2002] *MLR* 611.

⁵ D. Cooper *Sexing the City* (London: Rivers Oram Press, 1994); 'Local Government and Legal Consciousness in the Shadow of Juridification' [1995] *Journal of Law and Society* 506; 'Defiance and Non-Compliance: Religious Education and the Implementation Problem' [1995] *Current Legal Problems* 253; 'Institutional Illegality and Disobedience: Local Government Narratives' [1996] *Oxford Journal of Legal Studies* 256; 'Talmudic Territory? Space, Law and Modernist Discourse' [1996] *Journal of Law and Society* 529; *Governing Out of Order* (London: Rivers Oram Press, 1998).

challenged by citizens who seek to test the 'boundaries of authority' with the consequence that local authorities govern 'out of order'.

In terms of theory the works of Loughlin, Vincent-Jones and Cooper are diverse, in that each presents an analysis of a particular area of decision making, yet none of the works are irrelevant, unacceptable or even incorrect in terms of wider context of legal analysis on local government and the central-local relationship. In other words, the analyses are just very different, possibly even incompatible to the extent that it could be argued that that apart from writing upon the same topic, the only other feature which they share is that of theory.

Theory and Explanation Within The Context of Local Government And The Central-Local Relationship

The assertion that a shared feature of all the theorists is that of theory would appear to be somewhat simplistic. However, it is argued that theory, as a shared feature, relates to the use of theory as a mechanism for the explanation of knowledge. In this context theory can be seen to perform a dual role; firstly as a form of explanation and secondly, as a structure for explanation. The difference between theory as a form or structure is ontological, that is, in the nature of the explanation which is being sought. It is proposed to explore this difference in terms of legal analysis on local government and the central-local relationship in order to identify how theory is used and how the works of the various theorists may be connected through theory.

a. Theory As A Form of Explanation

Theory, as a form of explanation focuses upon the phenomenon to be understood. The goal of the analysis is to propose an explanation concerning a particular phenomenon without offering a specific view about the nature of the structure used to present the explanation. In the context of legal analysis on local government and the central-local relationship, this use of theory can be found to operate on a number of levels.

For example, a study which focuses upon the provision of a particular service within the context of local government may employ a specific theory as a means of explanation but ultimately the outcome of the study is a micro level analysis on the provision of the service. Explanation may be given in respect of the choice of theory but this done with a view to enhancing the validity of the study of the particular service rather than offering a critique of the chosen theory. At the meso level a study which, for example, seeks to focus upon the use of law within the central-local relationship⁶ may result in the construction of a particular theory. However, as a form of explanation the focus is the constructed theory and not the nature of theory. Some consideration may be given to the nature of the constructed theory but only as a means of justifying its validity.⁷ Of course, the constructed theory may, in turn, become the subject of further testing in order to ascertain how

⁶ Such as Loughlin. See n.3 above.

⁷ Such as Loughlin's desire to create a normative theory. See below.

accurate the theory actually is.⁸ The features of the theory may be examined but again the outcome is a study on the nature of the particular theory concerned rather than the nature of theory per se. Ultimately, a limitation of micro and meso level analysis is their contextual nature and this may make consideration of the nature of theory difficult and, it could even be argued, irrelevant.

Within legal analysis on local government and the central-local relationship micro and meso level analysis predominates. Certainly the works of Loughlin, Vincent-Jones and Cooper fall within the meso level category. However, the absence of a macro or meta-level analysis should not preclude its consideration, particularly as meta-explanation may provide an able mechanism for connecting the works of the theorists. Accordingly, a number of options can be considered. The theory of legal pluralism, for example, seeks to provide an explanation as to the nature and the authority of law. These are crucial issues within law in general, and jurisprudence in particular and it would appear that such an 'open textured' theory should be able to accommodate and explain a diverse range of analyses which the works of the three theorists represents. However, as a form of meta-explanation legal pluralism possesses limitations. For example, whilst legal pluralism may accommodate the work of Cooper, who seeks to include a diverse source of norms, Loughlin explicitly seeks to exclude non-official sources of norms and therefore cannot be easily incorporated. In other words, meta-theory as a form of explanation is only effective if the parameters contained within the theory to be accommodated allows for inclusion. Alternatively, the theory of constitutional pluralism could, for example be used to create a common connection. All three theorists share a specific parameter, that of local government, which does possess a constitutional dimension. Except that in this instance the approach represents a form of unification, as opposed to a form of connection. The meeting of theory occurs by way of lifting the features of the lesser theories (those of Loughlin, Vincent-Jones and Cooper) in order to converge with the reductive elements contained within the dominant theory (that of constitutional pluralism). In other words, any link between the three theorists will occur within the context of the theory of constitutional pluralism. Furthermore, since the link will be hierarchical in nature the possibility of a direct linear link between the works of Loughlin, Vincent-Jones and Cooper may be remote, difficult or even become overwhelmed by the theory of constitutional pluralism. Ultimately, the ensuing analysis may not even offer a form of connection for these theories, but result in the creation an enhanced theory of constitutional pluralism. Such an approach may possess merit if the aim is to produce a meta-analysis on local government in terms of the theory of constitutional pluralism. However, since the aim of this paper is to identify a means of connecting theory within the context of local government and existing legal analysis on this topic, the use of the theory of constitutional pluralism would appear to be unwarranted and possibly even divergent.

Ultimately, theory as a form of explanation possesses merit but since its function is to provide reasons or to demonstrate why and how a phenomenon

⁸ See B. Mauthe, *Walls of Illusion: Decision Making and Compulsory Competitive Tendering in Local Government* unpublished PhD thesis; below n.20 (1999).

has occurred, its role a method for connecting theories appears to be limited even when the context for analysis is shared. It could even be argued that theory as form emphasizes, rather than diminishes, the incompatibility of the work of the various theorists.

b. Theory As Explanatory Structure

Theory as an explanatory structure concerns the nature of theory as a mechanism for the explanation of knowledge. It is the nature of theory per se which is the primary focus of analysis rather than the nature a particular theory (or theories) or context. It is about the effect that theory as a structure can have on an explanation rather than how theory can be used as a mechanism for explanation.

Legal analysis on local government and the central-local relationship generally focuses upon theory as a form of explanation and it is from this perspective the works of Loughlin, Vincent-Jones and Cooper are generally presented and evaluated. Given the paucity of analysis within public law regarding the use of theory as an explanatory structure, it is proposed to consider further the works of Loughlin, Vincent-Jones and Cooper in terms of this approach.

Theory as Explanatory Structure and Legal Analysis on Local Government and the Central-Local Relationship

Theory as an explanatory structure presumes that there are certain features which theory, as a mechanism for the explanation of knowledge, will possess. It can, for example, be argued that key questions which are addressed in respect of every theory are those of how and why the theories have been produced. In other words, perhaps by focusing on features such as methodology, perspective and justification, it may be possible to identify a connection between the works of Loughlin, Vincent-Jones and Cooper.

a. Methodology

Methodology can be broken down into two categories, those of sources and choice.

(i) Sources

The sources used by the theorists are varied. Loughlin uses a non-empirical basis in that the substance of the analysis is legal rules, such as case law and statute. Loughlin also seeks to specifically exclude non-legal rules, such as administrative rules. The justification for this exclusion is that legal rules are the outcome of constitutional processes where procedures for the creation of rules and the control of rule based decision making are identifiable and represent a direct response to the values which operate within the political arena. By contrast, the norms and values which underpin the development of administrative rules originate within the administrative decision making system and are the product of informal, unaccountable professional

networks. Accordingly, not only do legal rules represent 'reality,'⁹ they possess legitimacy,¹⁰ unlike the administrative rules.

Vincent-Jones employs an empirical methodology in that a detailed examination was conducted in respect of the CCT transactions developed in a large metropolitan city council in the north of England.¹¹ The focus was upon the practical impact of CCT and the quasi-contractual mechanisms which emerged. In some respects approach is similar to Loughlin's in that there is a practical basis from which the analysis stems. However, it is possible to argue that the objectivity of Vincent-Jones base is more tangible in that it is based on actual observed practices whereas Loughlin focuses upon a series of identifiable events¹² which lead to outcomes within defined processes, such as, legislative and judicial decision making. These outcomes are then represented as observed facts.

The methodology deployed by Cooper is mixed in that it is part document based and part empirical. However, whereas for Loughlin and Vincent-Jones the purpose of methodology is to provide an explanation of a particular event, both actual and perceived, from which an understanding is constructed, for Cooper methodology is about the search for contingencies. That is, events or occurrences which can influence the present and the future. A suggested consequence of this difference is that whilst the sources used by Loughlin and Vincent-Jones represent certainty, Coopers represents uncertainty and consequently could be perceived to lack rationality.

(ii) Choice

Choice relates to the material which will, or will not, be included within the parameters of the particular study. Such decisions are generally represented as being 'theory neutral' but, it is suggested, the choices which the theorists make have implications of a theoretical nature.

As part of the construction of the theory of juridification Loughlin adopts concepts developed in the works of earlier theorists writing on local government, such as Elliot¹³ and Grant.¹⁴ Elliot introduced the concept of 'hierarchy' into legal analysis of the central-local government relationship which he represented as the differing constitutional status of local government and central government. Grant devised the concept of a 'central-local axis' which incorporates elements of the political, legal and administrative relationship that exists between the two tiers of government.

⁹ Above n.3 (1989).

¹⁰ The legitimacy is a consequence of the fulfilment of certain criteria and values to be found in the process by which the rule emerges. In the context of the theory of juridification of the central-local government relationship, Loughlin defines both the criteria and the values, which inevitably results in only legal rules possessing legitimacy.

¹¹ Above n.4 (1994), 215.

¹² Such as the election of the Conservative government in 1979, the 'market' philosophy of the new Conservative government and the increase in legislation and case law relating to local government activities.

¹³ M. Elliot, *The Role of Law in Central-Local Relations* (SSRC, 1981).

¹⁴ M. Grant, 'Central-Local Relations: The Balance of Power' in J. Jowell and D. Oliver (eds.), *The Changing Constitution* (Oxford: Oxford University Press, 2nd ed., 1989), 247.

Both these concepts imply that the central local-relationship is multi-dimensional. Hierarchy, for example, suggests that there are degrees of control whilst the central-local axis indicates that the relationship possesses a linear element in respect of the weight to be accorded to different forms of decision making. Loughlin however chooses to exclude these variances and focus exclusively upon legal decision making which he represents as occupying a superior, weighted and most importantly, a singular position in respect of the dimensions of hierarchy and the linear axis. How Loughlin is able to achieve this partial selection of knowledge is by focusing on those aspects of the concepts which accord with the parameters identified as forming the sources from which the theory of juridification will be based. Loughlin is then able to select Elliot's normative analysis on hierarchy, but exclude the descriptive constitutional principles which Elliot identified within the central-local framework. Furthermore, the descriptive features which Loughlin desires to include, that of legal decision making, are relabelled as normative.

The choices made by Vincent-Jones in terms of theory construction are different from that of Loughlin. Whilst Loughlin, from the outset, states that the concern is the wider picture of the central-local relationship Vincent-Jones initial analysis focuses upon a particular area of local government decision making. From this pragmatic level the study is eventually 'lifted' to consider the wider picture. However, the leap from CCT practices to the wider 'picture' of the central-local relationship occurs in two stages. The first stage entails altering the context of the study from CCT to the reconciliation of the operation of a private law mechanism, contract law, in a public law context, local government. The outcome of this stage is the construction of a model of local government which is represented as 'contractual governance'. From this the analysis is then lifted to consider the wider picture of the central-local relationship. What is particularly interesting about this incremental process is the 'lifting' mechanisms employed by Vincent-Jones. In shifting context, for example, concepts are used whilst in shifting level, theory is brought into play. Inherent within the process is the presumption that the knowledge found is universal.

Cooper seeks to identify a network of relations and causality over spatial and temporal dimensions.¹⁵ In terms of Loughlin's and Vincent-Jones this approach can be articulated as Loughlin focusing upon the spatial relationship of central/local government whilst Vincent-Jones focuses upon the temporal development of CCT practices. Cooper seeks to abandon these dimensions because they limit the story that can be told. Ironically Cooper, like Loughlin and Vincent-Jones, also focuses upon the exercise of power and control, but represents this 'governmental excess'¹⁶ as a testing of the boundaries of the state rather than a conflict concerning the possession and control of power. Cooper achieves this by examining how 'excess' is articulated by groups through activities that exist at the margins of society by virtue of features, such as religion, sexuality or animal welfare. The hunting of animals, for example, would not normally be seen as part of the state but

¹⁵ M. Dean *Critical and Effective Histories: Foucault's Methods and Historical Sociology* (London: Routledge, 1994), 93-94.

¹⁶ Above n.5 (1998), 4.

the hunt can be seen as part of the state structure when groups use it as a device to link a variety of processes, norms, and practices. This does not mean that Coopers methodological approach is that of 'anything goes'. As Cooper explains there are limitations, but these limitations exist in terms of the sources used, as opposed to the constructivist tools, such as, concepts and theory employed by Loughlin and Vincent-Jones.

The choices exercised by the theorist in terms of construction are quite varied. It could be argued that Loughlin opts for utility, in that the redefined concepts ultimately enhance the weight of the theory of juridification. Vincent-Jones incremental approach represents certainty. The basis for the study is that of accuracy and each subsequent stage of development can be traced back to this basis. Furthermore, the weight attached to each stage of development is enhanced due to the accuracy which underpins the whole analysis. Coopers choice is quite different in that it is more important to include as wide a range of knowledge as is possible, legal and non-legal, although there are rational limitations. In other words, fruitfulness is the guiding principle.

b. Perspective

Each of the theorists draws upon a non-legal discipline. Loughlin's analysis draws upon political theory, in particular, functionalism which is described as a critical examination of the 'reasoning process of the courts in order to expose the value assumptions on which they rest'.¹⁷ Loughlin argues that 'traditional legal theory' is not "indicative of the reality of relations within the administrative-political system" and its relationship with law.¹⁸

Whilst Loughlin argues that the functionalist style facilitates the development of legal analysis, it is possible to argue that the style is also capable of the converse, controlling the development of legal analysis. In the context of Loughlin's analysis this 'control'¹⁹ can be identified as assuming three distinct forms, that of limiting, directing and structuring. Limiting occurs by way of restricting the connection of law exclusively to political decision making. The justification for this is the representation of law as the product of specific historical political events and the desire to create a realistic model. Loughlin is dismissive of alternative forms of decision making, most notably administrative decision making. Yet in the context of the central-local relationship administrative decision making has been shown to have impact in terms of the ability of local authorities to make rules, even when the ability to make rules has supposedly been removed through juridifying policies such as CCT.²⁰ The problem for administrative decision making is that by its nature it is difficult to represent as the master of a

¹⁷ M. Loughlin, *Public Law and Political Theory*, (Oxford: Clarendon Press, 1992), 172.

¹⁸ Above n.3 (1989), 25.

¹⁹ See D. Feldman, 'Judicial Review: A Way of Controlling Government', *Public Administration*, Vol. 66, Spring 1988, p.21.

²⁰ See B. Mauthe 'District Auditors and Decision Making in the Central-Local Government Relationship' [1999] *Anglo-American Law Review* 447; B. Mauthe, 'The Notion of Rules and Rule-Making in the Central-Local Government Relationship' [2000] *Anglo-American Law Review* 315.

process of social evolution since it is generally perceived to be the facilitator of that process.²¹

The functionalist style also enables Loughlin to direct the form of law's connection with politics in that the connection is in respect of political decision making rather than the works of political scientists on concepts such as the state or power or even political theory, its application or development. Functionalism as a theory is not deployed, only the functionalist style, since to apply the theory would require explanation in order to meet the conditions of the theory. In other words, the connection is with the practical rather than the theoretical. The functionalist approach is used as an investigative tool to overcome the evaluative failings of legal analysis.

Finally, structuring enables Loughlin to make explicit the goals and values to be deployed in constructing the model. For example, legal analysis is first directed towards political decision making and informed that such decision making occurs prior to legal decision making. This allows legal analysis to ignore the actual details of political decision making, deploying it instead as descriptive explanatory background. It also confines legal analysis to proper constitutional principles, such as the status of local government as a decision making body within the wider institutional structure of the state.

The connection by legal analysis with political theory is complex. In some respects Loughlin has adopted a political theory yet manages to deploy it in a non-theoretical manner in order to construct a legal theory. Ultimately, it could be argued that it does not matter what form of model Loughlin produces, which methodology is used, whether the model is 'realistic'/'correct' or what form or manner of political theory is deployed. What Loughlin does achieve is to facilitate a wider framework for legal analysis of the constitutional lacuna which the central-local relationship was perceived to represent.

Whilst Loughlin is explicit in seeking to link law with a non-legal discipline, such a link is not so expressive within Vincent-Jones analysis. Certainly, the first impression of Vincent-Jones analysis is that it is firmly entrenched in law. Not only is there an examination of a legal mechanisms, CCT practices, the evaluation occurs against a coherent body of law, namely contract law. In the subsequent development of the model of 'contractual governance', this too is evaluated against the legal theory of autopoiesis. However, there is a link with a non-legal discipline, namely economics, although this connection is rooted deep within the subject matter of the analysis, the emerging CCT practices in local authority service provision. However, the dilemma facing Vincent-Jones is that the 'market' for local authority services is not a true market but a quasi-market. Accordingly, legal analysis which draws upon neoclassical economics, such as pure contract law, cannot apply, nor can the 'new' economic approach to law which focuses upon non-market law,²² such as the family.²³ However, the existence of the quasi-market, does facilitate

²¹ H. Simon *Administrative Behaviour* (New York: The Free Press, 4th ed., 1997).

²² P. Burrows and C. E. Veljanovski (eds), *The Economic Approach to Law* (London: Butterworths, 1981) 2.

²³ G. S. Becker *The Economic Approach to Human Behaviour* (Chicago: University of Chicago Press, 1977).

the deployment of legal analysis which focuses on the regulation of economic systems and the legal forms which maintain them, in other words, systems analysis and in particular, autopoiesis. Systems are self regulating mechanisms which seek to perpetuate their own existence whilst autopoietic systems take the notion of perpetuation further into that of self-reproduction.²⁴ As an autopoietic system law is self-reproducing in that it ultimately produces new legal acts.

The attraction of autopoiesis for Vincent-Jones is this feature of self-reproduction in that it facilitates explanation of the impact of CCT on local authority decision-making. CCT caused local authorities to alter their internal organisation and decision-making procedures. There occurred a fundamental shift 'from vertically integrated or hierarchical organisations' where social order and expectations were maintained through centralised command, bureaucratic administrative procedures to that of markets which require a separate institutional ordering and normative guarantees against risks and uncertainties that naturally arise in competitive contexts where the self-interest of business operators conflict'.²⁵ The self-reproduction occurs through local authorities 'neutralising' the legislation by adopting strategies of minimal compliance which left their core values intact.

A further attraction of autopoietic systems is their closed and self-referencing nature. An economic system, for example, will only recognise economic norms and not legal norms. Legal norms are 'external noise' which the economic system will filter and reconstruct in accordance with its own rationality of efficiency.²⁶ In other words, law works indirectly through 'reflexion structures' that encourage within the systems a self-reflection on the norms and values that should be guiding decision-making. In the context of CCT, local authorities self-reflection resulted in the adoption of 'the logic of competition, accounting separateness and a quasi-market organisation'.²⁷

There is merit to Vincent-Jones analysis. In terms of autopoiesis, Vincent-Jones is able to affirm much of its notions and, ironically, address some of its main criticisms. Autopoiesis, for example, has been criticised for its non-empirical base and the absence of a 'regulatory crisis'. Vincent-Jones analysis is an empirical study of a 'crisis' period in local government. A further criticism of autopoiesis is its tendency towards 'grand theory' – a criticism that can also be levelled at Loughlin. However, whilst this flaw is not evident in Vincent-Jones analysis the potential is there in terms of the methodology in that Vincent-Jones predilection to lift the analysis in stages could, if taken to the ultimate, result in the creation of a grand theory of local authority governance.

A final criticism is that of the intricacy of the analysis which occurs on two levels. Firstly, in terms of the analysis itself, it could be argued that the outcome of Vincent-Jones analysis is an evaluation of the responsive and

²⁴ See N. Luhman, 'The self-reproduction of the Law and its limits' in G. Teubner (ed.), *Dilemmas of Law in the Welfare State* (Berlin: Walter de Gruyter, 1985) and G. Teubner, *Law as an Autopoietic System* (London: Blackwell, 1993).

²⁵ Above n.4 (1998), 374.

²⁶ J. Black, 'Constitutionalising Self-Regulation' [1996] *Modern Law Review* 24.

²⁷ Above n.2 (1998), 373. See also note 8 above for an alternative interpretation on the impact of CCT.

reflective nature of law and the development of legal theory rather than an evaluation of the substantive law either in the context of contract or local government. In other words, Vincent-Jones analysis does not remain within the social framework from which it originated but moves into a jurisprudential context and therefore becomes removed from the reality which it sought to explain.

Secondly, there is complexity in the terms of the theoretical lineage of the analysis. Whilst it was possible to identify the route which has led to the construction of Vincent-Jones analysis, unravelling the route of construction of the concepts and theories that he deploys reveals an intensely complex web. Explained simplistically, systems analysis, for example, was initially developed in engineering and physics. It was then adopted by social scientists, ultimately discredited and then abandoned although it has been deployed by law through the theory of autopoiesis. Autopoiesis itself was developed in biology and adopted by 'sociology of law' thinkers in order to explain phenomena occurring in regulatory analysis²⁸ although the existence of the phenomenon, i.e. a crisis, is a matter of debate. Regulation is fundamentally a political-economic concept²⁹ and systems theory is just one approach, albeit one that is increasingly used for the analysis of law.³⁰ At each stage of development, theories will have been constructed to solve a particular problem. If knowledge of that particular problem is not passed to the next stage of development then all the values, norms, definitions, flaws, etc that went into the construction will pass into the new development unrecognised. New theories 'envelop' the old theories but the concepts which link one theory to another may not always demonstrate continuity but discontinuity.³¹

Accordingly, given the intricacies of the theoretical development, the link between Vincent-Jones analysis and economics could be described as tenuous, but it does exist. Vincent-Jones never sought to explain an economic reality but to identify why law behaves the way it does in a particular economic situation. In other words, the economic situation, just as Loughlin's political situation, is the catalyst for an analysis on the nature of law.

The non-legal discipline which Cooper draws upon is that of sociology. In particular, Cooper draws on Foucault's work of governmentality.³² The notion of governmentality is counterpoised to statist conceptions of power which Foucault regards as dominating modern analysis of social relations.

²⁸ See H. Rottleuthner 'Biological Metaphors in Legal Thought' in G. Teubner (ed.) *Autopoietic Law: A New Approach to Law and Society* (1988) p.97; M. King 'The Truth about Autopoiesis' [1993] *Journal of Law and Society*, Vol. 20, p.218.

²⁹ A. Ogus, *Regulation* (Oxford: Clarendon Press, 1994), 1.

³⁰ See B. Lange, 'Understanding Regulatory Law: Empirical Versus Systems-theoretical Approaches', [1998] *Oxford Journal of Legal Studies* 449.

³¹ See M. Tiles, *Bachelard, Science and Objectivity* (Cambridge: Cambridge University Press, 1984).

³² See G. Burchell, C. Gordon and P. Miller (eds.), *The Foucault Effect: Studies in Governmentality* (Chicago: Chicago University Press, 1991); A. Hunt and G. Wickham, *Foucault and Law* (London: Pluto Press, 1994) and A. Barry, T. Osborne and N. Rose (eds.), *Foucault and Political Reason* (Chicago: University of Chicago Press, 1996).

Foucault argues that to represent power as something which is institutional or prohibitive, does not explain the diverse and infinitesimal power relations that permeates all aspects of society. In other words, to focus exclusively upon law does not represent reality, but to identify reality it is necessary to acknowledge that law, politics, economics, etc are part of a complex web of power. Loughlin and Vincent-Jones would agree with this argument but differ from Cooper in terms of their identification of how the various powers are arranged. For Loughlin, law and politics are juxtaposed but politics presupposes law. A consequence of this presupposition is that laws, which are the product of political activity (which is its self established in the form of law), will possess legal validity and yet be free of extra-legal connections, such as politics.³³ Vincent-Jones takes this argument further in that politics is excluded even when laws are used to pursue specific ideological goals, such as, the use of markets. Cooper challenges the notion that law is a normatively closed tool for government policy. Law, like politics, is a technique of state power and can be subject to 'local interpretation and responses' which in turn are driven by ideological goals. In other words, law can never exist separate from politics even once laws come into existence. Even the interpretation of law and the response of institutions to law represents a form of political activity.

This desire for the inclusion of a diversity of power relations is manifested within Cooper's analysis in two ways. Firstly, Cooper goes beyond the traditional bi-institutional perspective of central government and local government to include civil bodies that speak about religion, sexuality and animal rights. These bodies do not possess formal legal powers but they do influence decision making. However, because of the breadth of agencies that are used, this means that Cooper's analysis must always remain at the meso level. It can never be used to evaluate decision making at the macro level, such as Loughlin, or even to form the basis of construction of a macro level model, such as Vincent-Jones. Whilst these wider/higher forms of analysis are not part of Cooper's goal, as a limitation it could be argued that such a limitation does not represent a flaw, since Cooper's aim is exposure rather than explanation or evaluation. Alternatively, as a limitation it may be perceived as a form of underdetermination in that once reality is exposed Cooper leaves nothing to provide for further development. Of course, this argument assumes that development, or the need to create further theory or to facilitate the further creation of theory, is the desire of theory and theorists. The approach also asserts, albeit covertly, that any alternative theoretical explanations are 'unreal'.

Secondly, Cooper's approach allows for the inclusion of 'narrative'. Narrative is the use of biographical information, for example, where someone tells of their involvement or perspective concerning a particular incident or conflict.³⁴ This verbal response of individuals is used by Cooper to demonstrate the interaction of law with people, politics, history, geography, identity, norms etc. It is an approach that neither Loughlin nor Vincent-Jones could deploy although Vincent-Jones will have acquired

³³ The same argument could be applied to Vincent-Jones but with the notion of politics replaced with that of the market place.

³⁴ Generally see A. Cavarero *Relating Narratives* (London: Routledge, 2000).

‘biographical information’ as part of the empirical base. The information is not used because within the context of methodology deployed such subject matter comes to represent objective observations. However, for Cooper, biographical information represents part of the text of the conflict. Again, a potential criticism of the use of narrative is that it limits analysis to the micro level.

c. Justification

Justification relates to the reasons that the various theorists present in terms of sources and choices. Loughlin, for example, specifically wishes to construct a model of the central-local relationship that is ‘realistic’.³⁵ Furthermore, this realism is to be achieved by adopting as the ‘principal focus the examination of the manner in which the normative structures of law can contribute to the tasks of guidance, control and evaluation in government’.³⁶ However, in creating this realistic and normative model, Loughlin begins from a descriptive premise, the evaluative failings of law, and then translates this into a normative premise, of how central government has employed its superior legal status in a relationship where law did not matter. Loughlin’s translations are possible because the concepts developed by Elliot and Grant have become intuitive in that they are accepted as *a priori* truths within legal analysis. The irony of this path of construction is that the reality and the objectivity of the model is not dependant upon Loughlin’s discovery of the political use of law but the passivity and the subjectivity³⁷ of the concepts.

It could also be argued that what Loughlin represents as ‘reality’ could in fact be represented as ‘correctness’. Reality represents ‘what is’, or fact, whilst ‘correctness’ is the product of procedures which result in a determinable end model –something which ‘ought to be’ rather than ‘what is’. Accordingly, it is possible to produce a model that is ‘correct’ but not necessarily ‘real’ since such an end model is a product of the chosen process of discovery rather than the actual discovered knowledge. It is also a method that validates the model since, like the method, it is reasoned and coherent. The outcome of such an approach is that any challenge to the model also represents a challenge to the method. This, in turn makes any challenge to the model problematical. After all, it is easier to ‘trash’ a ‘real’ model since all that is required is new evidence whereas to challenge a ‘correct’ model requires not just evidence but the identification as to why and what is flawed in the model. Loughlin’s model has been shown to contain flaws³⁸ but it’s ‘correctness’ is so entrenched within Public Law that there have been few direct challenges.

Vincent-Jones also wishes to provide an explanation of reality but, it is suggested, the reality which Vincent-Jones identifies is ultimately a representation of practice and theory. This assertion can be found through

³⁵ Above n.3 (1989), 25.

³⁶ Above n.17 (1992), 264.

³⁷ The subjectivity arises because unlike Loughlin, neither Elliot nor Grant sought to construct a model based on a particular perspective. Elliot and Grant also include both normative and descriptive elements within their analysis.

³⁸ Above n.8 and n.20 (2000).

examination of the methodology employed. The analysis begins with two forms of knowledge, CCT practices and contract law. The CCT practices are empirical evidence which is represented as newly acquired factual knowledge. Contract law, however, represents a pre-formed body of knowledge which possesses specific features, such as, abstract elements³⁹ and conditions⁴⁰ which ultimately provide the framework for its existence as a body of knowledge. It is against this body of knowledge that the emerging CCT practises are tested and the outcome is the construction of a model which is represented as 'near-contractual governance'. This model then becomes the object of further testing against another body of knowledge, the theory of autopoiesis, which is more abstract, less conditioned and possesses a more limited framework than contract law. The outcome is yet a further model, 'contractual governance' which in turn possesses the potential for further testing. Testing occurs against yet another theory, that of governmentality, which is again a theory that is more abstract, less conditioned and possesses a more of a limited framework than the original premise for construction, that of contract law. The analysis proceeds through a series of stages which ultimately represents a hierarchy of constructs.

A potential problem with this incremental approach is that the relationship between the original approximation and the ultimate approximation could become remote, possibly misplaced or even lost.⁴¹ Furthermore, with each stage of development the potential for abstraction and generalisation increases (a consequence of the emerging gap between each approximation) and the cycle for development becomes endless. Ironically, the whole process is represented as rational and objective, yet there is irrationality, it lies in the first step, that is, the choice of foundation from which to begin the analysis.

In some respects Vincent-Jones methodology is more open-ended than Loughlin's. The potential for the construction of models is infinite in that there is no limit other than the availability of suitable theories. Loughlin's model is, by contrast, finite in that it does not invite, or leave open, an avenue for development or revision. Instead, the only revision available is to test the validity of the theory itself which can ultimately lead either to its affirmation or destruction.

Like Loughlin and Vincent-Jones, Cooper also seeks to provide an explanation of reality but through an interpretation, or a 'reading', of specific events on the basis that political conflict is 'fluid' and 'constantly evolving'. The focus is on method as 'history' rather than method as deployment of a particular perspective (Loughlin) or method as development of theory (Vincent-Jones). The approach is directly attributable to the theorist that Cooper draws upon, namely 'Foucauldian political analysis, discourse theory, socio-legal studies and cultural geography'.⁴² However, it is possible to argue that the contingencies which Cooper includes are so diverse that

³⁹ For example, the notion of consideration.

⁴⁰ Such as the rules or circumstances which must be met in order for a contract to exist.

⁴¹ Vincent-Jones does attempt to address this problem through the extensive use of footnotes.

⁴² Above n.5 (1998) Preface.

they are incapable of producing coherent analysis of a phenomenon. Except that coherence only matters if the desired outcome is to identify error or accuracy. Since Coopers aim is to identify an alternative perception of reality, diversity, because of its breath should only produce exposure. The key to this exposure lies in how the contingencies are structured, in other words, there must be some commonality. Furthermore, this commonality goes beyond the mere assertion that there is commonality or that the only limitation is that of 'sources'. In other words, in there is a rationality to Coopers methodology. For example, Cooper argues that reality can be drawn from projections of relationships but, for the relationships to be valid there must be filters. In other words, Cooper states what 'real history' is to be included, rather than what is to be excluded. Ultimately each 'real history' becomes a 'representation' of a relationship which then forms the basis for a formal structure which then becomes the encapsulation of the particular phenomena. Ironically, the outcome is much the same as that of Loughlin and Vincent-Jones, the creation of a model/theory except that the aim is exposure rather than explanation or description.

d. Summary

The above examination on how and why the theories of Loughlin, Vincent-Jones and Cooper have been constructed reveals few shared features from which to identify a basis for interconnection. It could also be argued that the examination highlights, even further, how mutually incompatible the works are. Given the difficulties in connecting the works, in terms of theory, either as form or structure, it may be that theory is not the most effective method for achieving a connection. In other words, it might be more effective to abandon theory and consider an alternative mechanism for the explanation of knowledge, such as concepts.

Concepts as a Mechanism for the Explanation of Knowledge

There are many definitions as to 'what' is a concept⁴³ although basically they can be described as an abstract representation which attempts to use words to portray reality.⁴⁴ The terms concepts and theory are occasionally used interchangeably but as structures for the explanation of knowledge they are quite distinct. Theory is, for example, informal in nature in that there are no particular rules regarding its construction. The above examination of the works of the theorists revealed great diversity and incompatibility which would certainly indicate that, within the context of legal analysis on local government and the central local relationship, particular rules regarding theory construction do not appear to exist. Furthermore, even the role of theory was found to be dissimilar. Cooper, for example, seeks to use theory as a form of exposure rather than as a method for explanation. Concepts, on the other hand, appear to be the converse of theories. As a mechanism for the explanation of knowledge, concepts can seem to be both formal and fixed. Consider, for example, the notion of juridification. If a study on local government and the central-local relationship uses the notion of juridification

⁴³ Generally see M. Weitz *Theories of Concepts* (London: Routledge, 1988).

⁴⁴ Generally see W. B. Gallie 'Essentially Contested Concepts' in M. Black (ed.) *The Importance of Language* (Englewood Cliffs, New Jersey: Prentice-Hall, 1962); P. Thagard *Conceptual Revolutions* (Princeton: Princeton University Press, 1992).

as a concept, then for the concept to be understood there need only be a reference to a definition, such as that proffered by Loughlin.⁴⁵ It would be unnecessary to offer any further explanation since the concept would always be understood by virtue of the referenced definition. However, if the study wishes to use the notion of juridification as a hypothesis, then as a hypothesis it would be necessary to identify what are the features of juridification deployed by Loughlin. These features would be explained and it would be argued that, whilst it may be thought that these processes have occurred, it must be established that they have indeed occurred. In other words, juridification cannot be assumed it must be explained, identified and then proved. Accordingly, the distinction between a theory and a concept lies with where knowledge can be found. In respect of concepts, knowledge is pre-defined, whilst for theory, knowledge must be discovered however certain it may be thought that the knowledge exists.

Concepts And Legal Analysis On Local Government And The Central-Local Relationship

The use of concepts within legal analysis on local government and the central-local relationship is not unusual but, it is suggested, that the matter of how concepts are used, as a method for explanation, remains unexplored. Accordingly, it is proposed to consider how the various theorists use concepts in order to identify a number of perceptions regarding them.

Firstly, it can be found that concepts are generally presented as being fixed and inflexible in terms of the knowledge that they represent. The use of the concept of juridification, as indicated above, would appear to affirm this representation. Yet consider the concept of law as employed by Loughlin and Cooper. For Loughlin, law specifically relates to the formal structures from which statute and case law are produced whilst Cooper seeks to include a wide range of norms; social, political and even religious – hence the ‘governing out of order’. This diversity in the understanding of the presentation of law as a concept would suggest that concepts are not in fact rigid or fixed structures but possess an element of flexibility or ‘openness’. In other words, the pre-defined knowledge which makes up the understanding of the concept can be altered according to the context in which the concept is used. This flexibility allows for the presentation of alternative understandings by theorists who appear to share little beyond that of the context within which they write.

Secondly, within legal analysis on local government and the central-local relationship theory is used as the primary mechanism for the presentation of knowledge. Concepts are deployed as an explanatory tool but they are generally represented as a secondary or subsidiary mechanism to that of theory. Consider for example, the concepts of legalization and restructuring which form part of Loughlin’s theory of juridification. Yet, the manner in which Vincent-Jones deploys concepts, as a means for altering the level of analysis, would indicate that concepts do not have to exist within the domain of theory. Concepts possess a flexibility, in that they can be part of the structure of theory, yet exist apart from theory and at the same time facilitate

⁴⁵ Above n.3 (1989).

theory. In other words, just like theory, concepts can exist as a stand alone mechanism for the presentation of knowledge, ultimately even facilitating the explanation of complex structures where diverse and disparate theories may exist.⁴⁶

Accordingly, given the difficulties in connecting the works of the theorists in terms of theory, it is proposed to explore whether concepts may be able to facilitate such a connection.

Concepts As A Mechanism For The Presentation Of Knowledge In Respect Of Legal Analysis On Local-Government And The Central-Local Relationship

If concepts are to form the basis for the connection of the work of the three theorists, then the questions arises, which concepts should be deployed? The most obvious choice would be to select those concepts already used by the theorists, such as law, politics or government. Except that in using these concepts there are the risk of transferring the predefined knowledge attached to these concepts. Consider, for example the diverse interpretation employed by Loughlin and Cooper in respect of the concept of law. Accordingly, it is proposed to draw upon a number of concepts which should provide scope, simplicity, consistency and accuracy.

Within the work of each of the three theorists it is possible to identify a concept which provides a distinctive representation of their analysis. It can be argued, for example, that Loughlin focuses on 'structure', Vincent-Jones on 'rules', and Cooper focuses upon 'ideology'. In other words, the connection between all three theorists lies, not through theory, but through the examination of these concepts. Whilst these concepts may possess specific representation within the context of a particular theorist but, as concepts, they also have common elements which can be found within the works of all the theorists. It is proposed to examine this commonality.

a. Structure

A notional definition of structure is that it relates to framework, the various components within the framework, how these components relate to one another and how things, in general, are organised within the framework. Each of the theorists contains aspects of 'structure' within their work. Furthermore, 'structure' can be identified in assuming numerous forms.

Loughlin focuses upon the institutional structure of central government and local government. Central to Loughlin's perception of institutional structure is that local government is constrained and cannot choose or even define its relationship with central government. Any defining of the institutional relationship must be done either by central government or the judiciary through the mechanism of law. The solution offered by Loughlin to this constraint is also structural in nature. Initially, Loughlin argued in favour of

⁴⁶ See B. Mauthe 'The Notion of Sovereignty and its Presentation within Public Law: a Critique on the use of Theory and Concepts'[2005] *Northern Ireland Law Quarterly* 63.

a written constitution.⁴⁷ This solution, whilst being the most ideal, was also recognised as being the most unlikely and ultimately Loughlin settled for the creation of a constitutional norm of ‘locality’.⁴⁸ Although the normative status of the level of protection has altered, the focus of protection remains the same, the institutional relationship of central government and local government.

The second facet of structure that can be identified within Loughlin’s analysis lies in respect of the relationship of law to politics. Loughlin argues that political decision making exists prior to legal decision making in that a political decision to implement a particular policy may result in legislation. However, once the political decision has been made, and law is produced, then the relationship between law and politics ceases. How this represents a form of structure is that politics and law represent separate disciplines yet Loughlin’s approach allows for the articulation of a relationship. However, there are hidden facets within Loughlin’s approach since it is only the activities within the realm of law that enables the structural relationship with politics to be identified and maintained. Furthermore, the sequential position occupied by law within this structure ultimately enables law to ‘ignore’ politics from a specified, but self determined position. It also allows law to be selective as to which political activities to focus upon. Whilst the approach facilitates the inclusion of issues which previously were not possible within legal analysis of the central-local relationship,⁴⁹ it also acts a restraint in that law is seen as both a facilitator and enabler, but the limitations of law, whether conceptual or practical, are never explored.⁵⁰

Vincent-Jones within his analysis focuses on a different structure from that of the central-local relationship. Central government, through the use of law, requires local authorities to enter into an alternative, non-public law framework, that of the marketplace. Whereas Loughlin would focus on how this framework emerged, Vincent-Jones is more interested in exploring how local authorities operate within the new framework, particularly in respect of their relationship with private contractors. Whilst Loughlin represented structure as hierarchical in nature, for Vincent-Jones, structure is linear. The focus is also that of decision making but decision making based on relationships of equality within the market place. As a consequence, the concept of ‘structure’ is not a dominant feature in terms of the content of Vincent-Jones’ work and it only becomes a recognizable facet when there is a change in the structure. For example, when the market place is created and when it is altered as a consequence of the switch from CCT to Best Value. Such a change in the structure affects all the parties, hence its linear perception, unlike the operation of hierarchy within Loughlin’s model. It could be argued that where structure does dominate Vincent-Jones’ analysis is in relation to methodology. The incremental approach represents a discrete structure in that there is interplay between theory and practice within a strictly defined framework.

⁴⁷ M. Loughlin *The Constitutional Status of Local Government* (Commission for Local Democracy, 1994).

⁴⁸ Above n.3 (1996).

⁴⁹ Such as the central-local relationship as a mechanism for distributive justice.

⁵⁰ Such as highlighting the limitations of law in the absence of a formal written constitutional framework.

Cooper offers a very different perception regarding structure from that of Loughlin and Vincent-Jones. Although no formal definition of the concept can be found, nor is it a dominant feature of the analysis, structure can be found within the content of the analysis through the use of attributes such as community and geography. Cooper offers a spatial perception of the concept in that structure is a 'stretched out' notion which consists of a configuration of relations that operate at different levels and different ways. In other words, structure is about space, in particular local space, and how things are organised and done within that space. Where the boundary of local space ends does not matter, nor is it even an issue, since it is the activities within the local space that represents structure rather than space per se. Whilst different, there is some similarity to Loughlin and Vincent-Jones in that they too focus on activities albeit in respect of institutions (Parliament, the Executive and the Judiciary) or a context (the market place). Such wide variations on the concept would suggest that structure can be represented as multi-dimensional in nature, almost three dimensional, but, it is only by combining the works of the three theorists that this perception of structure in the context of local government can be found.

b. Rules

None of the theorists none offer a specific definition as to what constitutes a rule. However it is possible to identify a number of features concerning rules.

For Vincent-Jones, rules lie at the core of the analysis. Vincent-Jones observations on rule making by local authorities within the CCT regime provide the basis for the construction of the contractual governance model. This, in turn, leads to the construction of the theory of responsabilization. In other words, rules serve a dual purpose, those of subject and method. The rule making practices of local authorities represent observed knowledge, in other words they are Vincent-Jones's subject matter. But the explanation of these observations follows a reasoned progression, or method. Initially the knowledge is used in the construction of models and latterly in the deployment of theory. Whilst the progression is not represented as following a specific 'rule bound', formula, its continuous application by Vincent-Jones suggests, at least, the existence of a convention for the construction of models and theories.

For Loughlin, rules also lie at the core of the analysis in that legal rules, such as statute and case law, form the basis for the theory of juridification. Furthermore, Loughlin's perception of legal rules is that they represent the only real form of 'control' within the hierarchical structure of central government and local government.

Cooper does not focus on rules per se but on conflicts which can be based not only on law but also upon symbolic and cultural norms, factors which are often dismissed by lawyers as being irrelevant. These conflicts are about the stretching of the boundaries of authority rather than producing a definitive perception of what constitutes a rule or law.

Again it is possible to identify from the three theorists different facets of the same concept. Vincent-Jones perception of rules is about one area of law, contract law, accommodating another area of law, CCT. There is no formal requirement for these two areas of law to accommodate one another, just that

the consequence would be the failure of the local authority market with the possibility of direct interference by central government. It is therefore in the self interest of local authorities to find mechanisms for accommodation. Such economic and political factors are not apparent within Vincent-Jones analysis as the rules which emerge are represented to be a consequence of the responsive and reflexive nature of law. Loughlin focuses upon the authority that law possesses as a consequence of the operation of politics. There is a presumption that the power and authority of law are synonymous and that rules are the outcome of command and legitimacy. Cooper focuses on conflicts where law is only one of many factors. No explanation is offered regarding the nature of law, nor are there presumptions about what law can do. Yet, authority, power and command, features of law, exists within Coopers analysis, they are just not the overtly formal mechanisms identified by Vincent-Jones and Loughlin. This would suggest that in the central-local relationship the concept of rules is not static but depends upon numerous factors, such as context and the relationships being evaluated.

c. Ideology

Cooper defines ideology as ‘the various frameworks of meaning which justifies, legitimise, promote and emerge out particular networks of social relations’.⁵¹ Furthermore, ideology is not singular and its role is to create discourse. In the context of governance, ideology goes beyond beliefs and norms to include methods and paradigms. Accordingly, even conflicts which appear to be non-ideological, such as the banning of hunting⁵² or the increasing of Christianity’s profile within schools⁵³, possess an ideological component, since these issues ultimately concern differences in values and objectives.

Within Loughlin’s analysis the process of juridification is a direct consequence of the election of a Conservative government in 1979 with a specific ideological remit concerning the role and function of the state. The legal conflicts which emerge between central government and local government in the 1980’s are a direct consequence of this change in ideology. To some extent, Loughlin is chronicling the change in ideology. For Vincent-Jones, ideology underpins how rules emerge within the context of local authority transacting in the market place. Here, it is the impact of the new ideology on rules and rules making that are being chronicled. The conflicts which Cooper discusses are about how far dominant ideologies extend and lesser or alternative ideologies can challenge, through law and other mechanisms, dominant ideologies. In other words, how far does ideology operate as a form of control, beyond law.

It is suggested that ideology is a significant concept within the central-local relationship. Furthermore it links in with structure, since structural issues can often be a consequence of ideology (the basis of Loughlin’s analysis) and rules are often the consequence of ideology (Vincent-Jones).

⁵¹ Above n.5 (1994), 7–8 and 151–5. For Loughlin see above n. 1 (1989).

⁵² Above n.5 (1998), Chapter 7.

⁵³ Above n.5 (1995) and (1998), Chapter 3.

Conclusion

The aim of this paper was to demonstrate that although legal analysis of local government and the central-local relationship is incompatible in terms of theory, it is possible to connect the disparate analysis through the use of concepts. In order to achieve this aim, the notion of theory was explored in terms of its usage as either a form of explanation or as a structure for the explanation of knowledge. Far from revealing a means to connect the works of Loughlin, Vincent-Jones and Cooper, beyond the context within which they write, this approach appeared to enhance the diversity of their analysis. Accordingly, it was concluded that theory was not an appropriate device, by which to connect the works, and that it may be more effective to consider the use of an alternative mechanism for the presentation of knowledge, such as that of concepts. The nature of concepts was then considered along the perceptions concerning their usage within legal analysis on local government and the central-local relationship. It was found that, as an explanatory mechanism, concepts possess an openness and flexibility and could even be used to explain complex structures. Three concepts were then identified by which to evaluate the works of the theorist, those of structure, rules and ideology. The resulting analysis was a form of meta-explanation of a particular period in decision making within local government and the central-local relationship.

It is suggested that the approach contained within this paper could be transferred to other areas of public law analysis where a number of disparate and diverse theories exist in respect of a particular phenomenon. The approach has certainly been applied, with some success, in the respect of that most difficult of phenomenon with public law, sovereignty.⁵⁴ Given the vast differences between the notion of sovereignty and the topic of local government and the central-local relationship, this would indicate that the approach possesses merit.

⁵⁴ Above n.46.

THE CHURCH OF ENGLAND: SOME HISTORICAL REFLECTIONS ON A CONSTITUTIONAL CONUNDRUM

Charlotte Smith, Law School, University of Reading

Introduction

In recent years the conundrum of the constitutional and legal position and treatment of the Church of England has frequently been brought to the fore. Reform of the House of Lords and of the blasphemy laws has raised it in the context of initiatives for constitutional reform. Similarly, the implementation of the Human Rights Act 1998, and in particular the developing jurisprudence on the definition of public authorities under s6, has raised questions for the judicial treatment of the Church of England and its institutions. That the Church of England represents a conundrum for modern policy-makers and judges has, it will be argued, been decisively demonstrated in the case of *Aston Cantlow v Wallbank*,¹ which concerned the status of Anglican parochial church councils under s6 of the Human Rights Act.

This article seeks to cast light on the conundrum, firstly by examining what Establishment means in practical terms, and then by exploring it from the perspective of Church history in the last two hundred years. It will examine the constitutional, political and social changes which challenged the existence of an unproblematic connection between Church, State and nation,² and Church reactions to those changes. In doing so it will identify the ambiguities created by Establishment, and will illustrate how the changes of the later nineteenth century, continued and developed down to the present time, have created problems for modern attempts to address the legal and constitutional status of the Church of England.

Having sought some lessons from history, this article will conclude by demonstrating the applicability of the historical debate relating to Establishment to modern difficulties relating to the treatment of the Church of England. In doing so it will also comment, albeit briefly, on some of the other reasons why modern judges and policy-makers find it so difficult to give adequate attention and effect to the impact of the Church's historical and continuing Establishment.

¹ [2003] UKHL 37 and [2001] EWCA Civ 713.

² On this point see generally R. Brent, *Liberal Anglican Politics: Whiggery, Religion, and Reform, 1830-1841* (Oxford: Clarendon Press, 1987); A. Burns, 'The Costs and Benefits of Establishment: Clergy-Discipline Legislation in Parliament, c. 1830-c. 1870,' *Parliamentary History* (2000); and F. Knight, *The Nineteenth Century Church and Society* (Cambridge University Press, 1995).

Defining Establishment

Like almost every aspect of the organic and evolving British constitution, Establishment is extremely difficult to define and quantify with any reasonable degree of precision. As Paul Avis remarks:

“The establishment is a given history, a complex tissue of customary and statutory arrangements deeply embedded in the way society conducts its most serious affairs. Establishment has been profoundly modified and reformed over the past 400 years and will go on evolving. No one particular concrete expression of it is sacrosanct.”³

In its original formulation, in the Reformation Statutes, it appears that it merely signified that the Church had been settled or stabilised following a period of political instability and change. Thus the Latin rendering of the words ‘by law established’ in the Canons was *legibus stabilitam*.⁴ Over time, however, Establishment has come to mean far more than this. At its most basic level it embodies a relationship between Church and State which is distinct from that enjoyed by other religious bodies. In the words of Lord Selborne:

“The “Establishment” of the Church of England consists in certain relations, different from those of other religious bodies, in which the Church of England now stands towards the State, or the Public Law and Government of the country.”⁵

The central features of State recognition and support of the Church, and some degree of State authority and influence in its affairs, typify this special relationship.⁶ It is, in its essence, an assertion that religion is, in part at least, a public matter with which the State may (should) concern itself. It asserts that the State has a duty to provide for the moral and spiritual welfare of its citizens and that the Church, supported and recognised by the State, is the means by which the State fulfils that duty.⁷ As such Establishment exists in two aspects. In part it refers to a relationship between the Church and the institutions of the State, that is, to Parliament and the Crown. Establishment also, however, refers to the relationship between the Church and the citizens of the State, that is, to the nation.

Seeking Lessons in History: Church and State in The Nineteenth Century and The Creation of The Conundrum

What follows is an attempt to trace relations between Church, State, and nation in the nineteenth century, and to identify the historical and theological

³ P. Avis, *Church, State and Establishment* (London: S.P.C.K., 2001) p.34.

⁴ *ibid.*, pp.18-19.

⁵ R. Palmer, *The Endowment and Establishment of the Church of England. An Address Delivered at Blackmoor on Monday, January 11, 1866* (London: Cassell and Co., 1886) p.3.

⁶ Although this State authority may be weakened beyond any significant meaning – as in the present Establishment of the Church of Scotland.

⁷ For a good overview of Anglican Church-State theory see generally P. Avis, *Church, State and Establishment*.

forces which, it will be argued, have contributed to current difficulties regarding the legal and constitutional treatment of the Church of England.

Establishment as a Relationship between the Church and the Organs of the State – The Legal and Constitutional Aspect

Almost from its birth the Church has enjoyed an intimate relationship with the Crown, reflected in the role of its bishops as royal advisers and councillors, and in their right to sit in the House of Lords as advisers, representatives of the people and the owners of significant tracts of land. At the Reformation this relationship came, by virtue of the need to exclude the foreign jurisdiction of the Pope, to focus, and to pivot upon, the Royal Supremacy. This was embodied in section 8 of the Act of Supremacy 1558,⁸ which affirmed that:

“And . . . that suche jurisdictiones privileges superiorities and preheminences spirituall and ecclesiasticall, as by any spirituall or ecclesiasticall power or auctorite hathe heretofore bene or may lawfully be exercised or used for the visitacion of the ecclesiastical state and persons, and for reformation order and correccion of the same and of all maner of errours scismes abuses offences contemptes and enormities, shall for ever by auctorite of this present Parliament be united and annexed to the imperiall crowne of this realme; . . .”

The first practical expression of this was found in the law regulating the appointment of bishops.⁹ Having excluded the jurisdiction of the Pope Henry VIII first asserted the right to make all Episcopal appointments in the Appointment of Bishops Act 1533.¹⁰ Under the process established by this Act the Dean and Chapter of the relevant see were granted a licence (the *congé d'élire*) to appoint a bishop under the Great Seal. However a letter missive, stating the name of the person to be appointed, accompanied the licence.

The second practical manifestation of this relationship was found in the role of the Crown in the promulgation of Anglican canon law. By the Act of Supremacy 1533, and the Act of Submission of the Clergy 1534,¹¹ the clergy recognised the Crown as their law-maker, and forfeited any right to legislate for Church matters independently of the Crown. Under s1 of the Submission of the Clergy Act the Crown's licence and assent were required before canons made in the Convocations of the Church could be promulgated. By section 3 of that Act it was declared that the Church could not pass canon law which was contrary to the royal prerogative, the common law, statute law or other customs of the land.

A third practical consequence of Establishment was the role of the Queen in Parliament in the government and general law-making of the Church. The general effect of the Act of the Submission of the Clergy 1534 was to vest in Parliament a wide authority in ecclesiastical matters. By that Act the

⁸ 1 Eliz. c.1.

⁹ And certain other senior officers of the Church.

¹⁰ 26 Henry 8 c.1.

¹¹ 26 Henry 8 c.1.

Convocations of the Church were unable to meet except at the behest of the Monarch. In fact these bodies were moribund for much of the period leading up to the 1850s, with the result that, in practice, the Church had for a long period no national body except Parliament in which to deal with its administration and government. Similarly, just as the Church had no power to pass canon law except by the licence and assent of the Crown, it had no power to pass ecclesiastical legislation binding upon both the clergy and the laity. In practice, if legislation was required, it had to rely on the intervention of Parliament.

A further consequence of Establishment was found in the character and status of Anglican ecclesiastical law. That body of law was heterogeneous in character. It was comprised of Canons passed before the Reformation,¹² Canons promulgated by royal licence after the Reformation,¹³ statutes,¹⁴ and things done pursuant to statutory or prerogative powers.¹⁵ However, whatever form that ecclesiastical law took it formed part of the ordinary law of the land and was ultimately enforceable as such. In this respect it was fundamentally different to the law of other religious organisations, a situation identified by many as being the essence of Establishment.¹⁶

The final result of Establishment was the legal position granted to the ecclesiastical courts of the Church of England. The first aspect of this was the coercive jurisdiction which the State conferred upon those courts. These were recognised as public courts whose judgments were enforced in the last resort by the coercive powers of the State. Further, the final appeal in ecclesiastical causes was to the Crown.¹⁷ This appellate jurisdiction was exercised by the Court of Delegates until 1832, when it was transferred to the Judicial Committee of the Privy Council by the Privy Council Appeals Act of that year.¹⁸

Anglican Theory and Ideology

As demonstrated, Establishment resulted in a close relationship between Church and State. This relationship has been the subject of much theorising in the centuries following the Reformation, reflecting the fact that Establishment consists not only of a legal framework, but also of a web of theory and sentiment which has grown up around the law.¹⁹

Historically one of the most prominent features of Establishment theory was that Establishment was founded upon the idea of a union or unity of Church and State. Just as the Royal Supremacy was seen at the foundation of the

¹² Henry VIII expressed his intention at the Supplication of the Ordinaries, 1532, to subject this body of law to a commission of review but this work was never done.

¹³ Such as the canons of 1603.

¹⁴ Such as the various Acts of Uniformity.

¹⁵ Such as the *Book of Common Prayer* of 1662 and the Articles of Religion.

¹⁶ For statements on this point see R. Palmer, *The Endowment and Establishment of the Church of England*, p.43 and the Acts disestablishing the Churches of Wales and Ireland.

¹⁷ See the Act in Restraint of Appeals 1533, 24 Henry 8 c.12.

¹⁸ 2 & 3 Will. 4 c.92.

¹⁹ See P. M. H. Bell, *Disestablishment in Ireland and Wales* (London: S.P.C.K., 1969) introduction and chapter one.

legal expressions of Establishment so, too, it was found at the heart of the union or unity of Church and State. Church and State were united under the Queen as temporal governor of both.

The essential unity of Church and State was expressed by many thinkers in the idea that Church and State were the same body – the nation – seen from different perspectives. This theory was best known in connection with the writing of divines such as Richard Hooker²⁰ and Edmund Burke.²¹ It was also famously expressed in the nineteenth century by, among others, Thomas Arnold and Samuel Taylor Coleridge. Thomas Arnold, for example, idealising the union of Church and State, and arguing against the fragmentation of life into secular and spiritual elements, wrote:

“The natural and true state of things then is, that this power and this wisdom [of Church and State] should be united; that human life should not be pulled to pieces between two claimants, each pretending to exercise control over it, not in some particular portion, but universally; that wisdom should be armed with power, power guided by wisdom; that the Christian Church should have no external force to thwart its beneficent purposes; that government should not be poisoned by its internal ignorance or wickedness, and thus advance the cause of God’s enemy, rather than perform the part of God’s vicegerent.”²²

Coleridge maintained a careful distinction between the nation - of which the Church was a part – and the State – of which the Church was not a part, but to which it was joined. He also distinguished between the national Church – which was joined in a relationship with the State – and the universal Church – which owned no temporal allegiances and looked to Christ alone as its authority. Despite making these careful distinctions he too was clearly concerned to demonstrate the essentially close relationship between Church and State. He instructed his readers in this using the following parable:

“As the olive tree is said in its growth to fertilise the surrounding soil, to invigorate the roots of the vines in its immediate neighbourhood, and to improve the strength and flavour of the wines; such is the relation of the Christian and the national Church. But as the olive is not the same plant with the vine, or with the elm or poplar, (that is, the State) with which the vine is wedded; and as the vine with its prop may exist, though in less perfection, without the olive, or previous to its implantation; even so is Christianity. . .”²³

In both quotations a further aspect of Establishment was drawn to the attention of the reader. This union or unity of Church and State, or perhaps of Church and nation, existed for a reason. At its heart was an assumption that the State, by virtue of Establishment, was in some way underpinned by

²⁰ (1554-1600). His most famous work is *Of the Laws of Ecclesiastical Polity*.

²¹ (1729-97).

²² T. Arnold, *Fragment on the Church* (2nd ed.) (London: B. Fellowes, 1845) p.11.

²³ S. T. Coleridge, *On the Constitution of Church and State According to the Idea of Each* (3rd ed.) (London: William Pickering, 1869) p.60.

the Christian faith of the Church to which it was joined, and that the mission of the Church was benefited thereby.²⁴ Thus, for both Coleridge and Arnold, the essential virtue of the union of Church and State was that it furthered the mission of those bodies to fight evil and to improve the moral state of man. In this idea was found another bulwark of the constitutional relationship of Church and State, that is, the idea that not only did the Church have a spiritual identity and authority but so, too, did the State.

One of the most obvious foundations of the spiritual identity and authority of the State was found in the Royal Supremacy. This articulated the belief that the Monarch, under whom Church and State were united, was a species of spiritual person, though not a priest, whose authority was derived from God.²⁵ Article XXXVII of the Articles of Religion expressed this in the following words:

“Where we attribute to the King’s Majesty the chief government . . . we give not to our princes the ministering either of God’s Word, or of the Sacraments. . . but that only prerogative, which we see to have been given always to all godly Princes in holy Scriptures by God himself; that is, that they should rule all estates and degrees admitted to their Charge by God, whether they be Ecclesiastical or Temporal, and restrain with the civil sword the stubborn and evildoers.”

Equally important to the spiritual identity of the Church, especially as Parliamentary Sovereignty assumed its place as the keystone of the British Constitution, was the identification of Parliament as a spiritual body. Hence its traditional identification as the ‘King/Queen in Parliament under God.’ Like the Monarch it acted under the authority of God. Of equal importance to the Church, perhaps, was the identification of Parliament as being, in matters of ecclesiastical legislation, a lay synod. This identity was maintained until the nineteenth century by legislation limiting membership of Parliament to men who conformed with the Church. Acts such as the Tests Acts of 1673²⁶ and 1678²⁷ and the Corporation Act of 1661²⁸ ensured that those participating in the political life of the nation were, at least as judged by their actions, in communion with the Church.

Anglican Theory and Ideology Challenged: Lessons from the Nineteenth Century

From the middle of the nineteenth century the evolution and reform of the British constitution required Establishment’s supporters to respond to novel challenges as the character of that constitution changed. At the most basic level these changes refuted categorically the assumption that the close relationship of Church and State was to be supported because it demonstrated that the State was founded upon Anglican religion. When the State repealed

²⁴ See P. Avis, *Church, State and Establishment*.

²⁵ For a more detailed discussion of this see I. M. Mackenzie, *God’s Order and Natural Law: The Work of the Laudian Divines* (Aldershot: Ashgate, 2002) chapter five.

²⁶ 25 Car. 2 c.2.

²⁷ 30 Car. 2 c.1.

²⁸ 14 Car. 2 c.1.

the Tests and Corporation Acts in 1828, passed the Roman Catholic Relief Act 1829,²⁹ and later the Jews Relief Act 1858,³⁰ it demonstrated that there had been a shift in favour of a State founded upon democratic and pluralist principles.

Despite the changes to the constitution embodied in the Reform Acts, and in the progressive removal of the civil disabilities of Protestant Dissenters, Roman Catholics, Jews and atheists, there was no generalised challenge to the role that Christianity played in the Constitution. Many thinkers continued to defend that role, and the continued Establishment of the Church of England, on the basis that the services of religion morally improved men, and that there was a link³¹ between religion, morality and the foundation of a democratic society ruled by law rather than force.³² Thus Lord Selborne, defending Establishment, announced his conviction that:

“The mainstay of all law . . . is . . . the moral sense of mankind; religion, wherever it is truly professed, is . . . powerful in the direction and reinforcement of that moral sense.”³³

Many supporters of Establishment also continued to defend Establishment on the basis that the Established Church was best able to provide for the spiritual needs and moral education of the nation, and that this was both necessary to – and a duty of – the democratic State.³⁴ Thus Gladstone argued that the alliance of Church and State was ‘calculated to extend and establish the vital influences of Christianity, and therewith to increase and purify the mass of human happiness.’³⁵ Similarly, Spencer Holland asked:

“Can the State when aware that masses of people will never seek any form of religion or moral training of their own accord, and of the inability of private resources to cope with the wants of great populations hurried and worried with much service,- can it wash its hands of all responsibility, and blandly reply, “Am I my brother’s keeper?””³⁶

Given this viewpoint many writers, even late into the nineteenth century, continued to identify the State as being in some sense a spiritual or divinely ordained body. Lord Selborne, for example, calmly asserted that:

“Both powers, the Ecclesiastical and the Civil, are... in their origin divine; each has its proper province, the persons over which each exercises its authority are to a great extent the

²⁹ 10 Geo. 4 c.7.

³⁰ 21 & 22 Vict. C. 49

³¹ Now rejected. See HL Paper 95-I, appendix three on religious offences.

³² See, e.g. S. T. Coleridge, *On the Constitution of Church and State*, p.58 and S. Holland, *The National Church of a Democratic State* (London: Rivingtons, 1886) p.5.

³³ R. Palmer, *A Defence of the Church of England* (5th ed.) (London: Macmillan & Co, 194), p.73.

³⁴ See e.g. Coleridge, *On the Constitution of Church and State*, p.78 and S. Holland, *The National Church of a Democratic State*, p.8.

³⁵ W. E. Gladstone, *The State in its Relations to the Church* (2 Volumes) (4th ed.) (London: John Murray, 1841) p.3.

³⁶ S. Holland, *The National Church of a Democratic State*, p.10.

same; and there is, by the unchangeable laws of nature and Providence, a middle or common ground between them.”³⁷

However two of the ideas underpinning the spiritual identity of the State were fundamentally undermined by the constitutional changes effected between 1689 and the nineteenth century. The first idea affected in this way was that of the spiritual identity of the Monarch. So threadbare was this seen to be in reality that, speaking to a gathering of Churchmen in 1882, one conservative Churchman was moved to dismiss it out of hand,³⁸ saying:

“The theory of an anointed king as being a sort of *mixta persona* – part secular and part spiritual – though of considerable age, I should hardly consider as worthy of serious discussion.”³⁹

There were four key features of Divine Right which helped to underpin that spiritual identity: the idea that the Monarchy was a divinely ordained institution, the indefeasibility of the hereditary right of the Monarch, the idea that the Monarch was accountable to God alone, and the principle that non-resistance and passive obedience were ordained by God.⁴⁰ Two of these features were swept aside by the Glorious Revolution. The interruption of the Succession in 1689 destroyed the practical reality of any argument for the indefeasibility of the hereditary right of the Monarch. The victory of the notion of Parliamentary Sovereignty and the constitutional or limited Monarchy, as embodied in the Bill of Rights 1689, was fatal to the principle of a Monarch answerable to God alone.

The identification of Parliament as the lay synod of the Church, another manifestation of the spiritual identity of the State, was also undermined. When, under the legal changes of the nineteenth century, Parliament admitted Protestant Nonconformists, Roman Catholics, Jews and even non-believers to its seats, its identity as a lay synod for the Church, or even a Christian legislature, was increasingly lost from sight.

When the State embraced pluralist principles, and when its spiritual identity was obscured, it was perhaps unavoidable that the traditional concept of the unity or union of Church and State should be called into question. Thomas Arnold, that most ardent supporter of the unity of Church and State as different aspects of the nation, gave a warning to this effect when he asserted that the nineteenth century Church was not as one with the State. While arguing that this situation had no logical or rational foundation in theory he nevertheless maintained that Church and State were in fact separate and out of step.⁴¹

This feeling that Church and State, though still unified or united by law were in fact and theory divided and increasingly distinct, was fostered by the

³⁷ R. Palmer, *A Defence of the Church of England against Disestablishment*, p.73. See also W. E. Gladstone, *The State in its Relations to the Church*, p.50.

³⁸ Although see I. Bradley, *God Save the Queen: The Spiritual Dimension of the Monarchy* (London: D.L.T., 2002) for a discussion of the continuing spiritual role or identity of the monarch.

³⁹ Earl Nelson, ‘Church and Crown,’ *Report of Church Congress* (1882) p.201.

⁴⁰ N. Figgis, *The Theory of the Divine Right of Kings* (Cambridge, 1896) pp.5-6.

⁴¹ T. Arnold, *Fragments on Church and State* (London: B. Fellowes, 1845), pp.57-9.

theological concerns of, among others, one prominent group of Anglican thinkers. The Oxford Movement, reacting to the constitutional change and uncertainty of its times, sought for the Church an identity and authority entirely separate from its relationship with the State. Its members found this in an emphasis upon the authority and identity of the Church of England as part of the Church of Christ. In doing so, however, they threw into stark relief a distinction between the earthly and man-made authority of the State, and the God-derived spiritual authority of the Church. In a characteristic statement, for example, one of their leaders cried:

“CHRIST has not left this Church without claim of its own upon the attention of men . . . I fear we have neglected the real ground upon which our authority is built, – our APOSTOLICAL DESCENT.

We have been born, not of blood, nor of the will of the flesh, nor of the will of man, but of God. The Lord JESUS CHRIST gave His Spirit to His Apostles; they in turn laid their hands upon those who would succeed them; and these again on others; and so the sacred gift has been handed down to our present Bishops, who have appointed us as their assistants, and in some sense representatives.”⁴²

Establishment as a Relationship between the Church and the Nation – The Aspiration and Manifestation of a National Church

The Church of England has historically aspired to the status and role of a national church. This was most strongly evident in its approach to three main areas of its life: the definition of its aim and mission; the definition of its community or membership; and its pastoral and administrative organisation.

Turning first to the mission of the Church of England as a national Church, the essential feature was that it defined that mission in relation to the nation as a whole. It aspired to serve all citizens living in the State. In the words of a nineteenth century theologian:

“That which is essential to a national Church is, that it should undertake to assist the spiritual progress of the nation and of the individuals of which it is composed, in their various states and stages. . .”⁴³

This identification of the mission of the Church as being a mission to the nation as a whole reflected the history of the Church as a manifestation of public provision for the religious needs and spiritual welfare of the nation.⁴⁴

⁴² J. H. Newman, ‘Thoughts on the Ministerial Commission,’ in *Tracts for the Times* (London: J. G. & F. Rivington, 1839) p.2.

⁴³ H. B. Wilson, ‘Seances Historiques de Geneve,’ p207 in *Essays and Reviews* (12th ed.) (London: Longmans, 1969) (1st ed., 1860). Wilson was condemned by Convocation for this essay but the point he made about the mission of the Church was uncontroversial.

⁴⁴ For a full elucidation of these ideas see the memorandum of A. Pearce to the Joint Committee on Human Rights – published as an appendix to its 7th report of the 2003-4 session (HL 39, HC 382).

This was further reflected in the Anglican approach to the definition of its community or membership.

Historically it was seen as wholly fallacious to talk of the membership of the Church of England as a body discrete from the nation as a whole.⁴⁵ The Church began its life as a national church at a time when membership of Church and State was broadly co-extensive. Throughout much of its post-Reformation history civil and ecclesiastical law co-operated to maintain this idea of co-extensive membership, acting ruthlessly to solve the problems posed by incipient dissent. Hence dissenters, falling into the category of schismatics, were denounced and rendered *ipso facto* excommunicate by canons 9 and 12 of 1603.⁴⁶ Excluded from the membership or community of the Church under canon law, dissenters were further excluded from the political life of the nation by the action of statute. Prior to the constitutional reforms of the nineteenth century Acts such as the Test Acts of 1673 and 1678,⁴⁷ and the Corporation Act of 1661⁴⁸ ensured that dissenters were excluded from the civic life of the nation.⁴⁹ Historically, then, the Church was able to adopt an involuntary model of membership. Citizens of the State did not have to take positive steps to become members of the national Church, rather, they were assumed to be members unless they took positive steps to set themselves outside of its community.

Even following the reforms of the nineteenth century, which broke the link between religious affiliation and participation in the full civic life of the nation, and even in the face of religious pluralism and diversity, the nationality-based and involuntary model of Church membership continued.⁵⁰ Generally speaking it was for any individual to determine the question of membership. If he claimed to be a member, and if he wished to avail himself of its services, then both the Church and the law vindicated his claim. In the words of Lord Selborne:

“It has sometimes been said that every Englishman is a member of the Church of England . . . I imagine that few Dissenters are in the habit of so describing themselves . . . It is true, now as much as ever, that the Church does not repel from the rights and privileges of Church membership any persons, baptized and not excommunicated, who honestly seek, or

⁴⁵ Although for various attempts to do so for legal purposes see, e.g. *General Assembly of the Free Church of Scotland v Lord Overtoun* [1904] AC 612-3, per Lord Halsbury; *Re Selby's Will Trusts* [1966] 1 W.L.R. 43 – per Buckley J; *Marshall v Graham* [1907] 2 KBD 112, at 124 per Lord Alverstone CJ; *Re Perry Almshouses* [1898] 1 Ch 400 – per Stirling J. It can be noted that these cases tend to determine when a claim of membership should not be rejected, rather than defining a test of membership *per se*.

⁴⁶ G. Bray (Ed.), *The Anglican Canons 1529-1947* (Boydell Press: Church of England Record Society, Vol. VI, 1998).

⁴⁷ 25 Car 2 c.2 and 30 Car 2 c.1 respectively.

⁴⁸ 14 Car 2 c.1.

⁴⁹ Although successive Indemnity Acts were passed after 1727, mitigating the effects of these statutes.

⁵⁰ Although see generally M. J. D. Roberts, *The Role of the Laity in the Church of England c1850 – 1885* (DPhil, Oxford 1974) for a discussion of the modifications to, and pressures upon, this model of membership.

willingly accept them. It is also true that the law does not, without proof of the fact, presume any man to be a Dissenter. But the question is one of fact. The law is not so unreasonable as to call the same man at the same time a Churchman and a Dissenter.”⁵¹

The identification of the Church as a national institution, comprehending and ministering to all within that nation who did not reject its services, was reflected in the traditional organisation and administration of the Church: the parochial and territorial systems. These were the means by which it supported that mission.

Under the parochial and territorial systems the infrastructure of the Church stretched across the geographical bounds of the State and was entrenched within its physical fabric. Dividing the country into progressively smaller geographical units, first into provinces, and then into dioceses, archdeaconries, rural deaneries and parishes, they provided a system of religion which served every region and every inhabitant of the land, provoking Coleridge to the following paean of praise:

“That to every parish throughout the kingdom there is transplanted a germ of civilization; that in the remotest villages there is a nucleus, round which the capabilities of the place may crystalize and brighten; a model sufficiently superior to excite, yet sufficiently near to encourage and facilitate, imitation, the unobtrusive, continuous agency of a Protestant Church Establishment, this it is which the patriot and the philanthropist, who would fain unite the love of peace with a faith in the progressive amelioration of mankind, cannot estimate at too high a price.”⁵²

The historic legal rights of the laity in respect of the Church were defined within this framework.

The first legal right of the laity was to attend the public services of the Church. As Cockburn LCJ asserted, in *R v The Bishop of Oxford*:

“The Church is upheld and exists for the spiritual benefit of the laity. It is the right of the latter to take part, under the ministrations of the clergy, in the public worship . . . according to the ritual of the Church as by law established.”⁵³

Every English man, woman and child, not being excommunicate, had a legal right to attend the public services of the national Church performed according to law. The expression of this right reflected the lineaments of the parochial system. In *Taylor v Timson*, for example, it was made clear, both that the incumbent of every parish had a duty to provide and perform the services of the Church according to law, and that there was a general right for an individual to attend the services of his *parish* church.⁵⁴ In other words

⁵¹ *A Defence of the Church of England Against Disestablishment*, pp.195–196

⁵² S. T. Coleridge, *On the Constitution of Church and State According to the Idea of Each*, p.78.

⁵³ (1878-79) 4 Q.B.D. 260.

⁵⁴ (1888) 20 Q.B.D. 671.

the rights of the laity in any particular church were determined by the geographical relationship between any particular layman and the church.

The legal right to attend the public services of the Church was not the only right of citizens in respect of the Church. Other rights included the right to have their children baptised there, the right to be buried in the churchyard and the right to be buried according to the rites of the Church of England.⁵⁵

The system of religious provision and legal rights just described was, until the later nineteenth century, supported by a system of public taxation in the form of church rates and tithes. In addition to these, and still persisting today, there were the historic endowments of the Church in the form of rectoral and glebe lands. Together these sources of funding, property rights and obligations were designed to provide for the upkeep of parish churches and the maintenance of the parochial clergy.⁵⁶ The rector of the parish was legally entitled to the full ecclesiastical dues of a parish under law and was consequently under a duty to keep the chancel of the church in a state of repair.⁵⁷ If the rector was a lay person, meaning that the cure of souls was placed in the hands of a vicar, then the incumbent was supported by receipt of the lesser tithes and glebe land.⁵⁸ Parishioners paid tithes and church rates and, in return for a right to a seat⁵⁹ in the parish church, were placed under a legal obligation to keep the nave of the church in a state of repair.⁶⁰

The National Church Challenged: Lessons from the Nineteenth Century

It would be wrong to suggest that there was any wholesale abandonment of the national ideal and the national mission of the Church of England in the nineteenth century. However, it is undeniable that the nation which the Church aspired to serve, and the constitutional framework within which it aspired to do so, changed radically in the mid to late nineteenth century. This was to have a profound impact upon the Church.

England in the late nineteenth century, though far from being the pluralist and increasingly secular State of the twenty-first century, was no longer, either in theory or reality, a Confessional State. Non-Anglicans had been accepted into the full life of the nation as citizens. Further, increasing

⁵⁵ *Aston Cantlow v Wallbank* [2003] UKHL 37, at 130 – though these rights are subject to certain exceptions.

⁵⁶ Limitations on space preclude a full discussion of this complex arrangement in its history. See Law Com. No. 152 (1985) *Property Law: Liability for Chancel Repairs* and J. H. Baker, 'Lay Rectors and Chancel Repairs', (1984) 100 LQR 181-185.

⁵⁷ W. Dale, *The Law of the Parish Church* (7th ed.) (London: Butterworths, 1998) p.27.

⁵⁸ *ibid.* The cure of souls could also be held by a curate who had no property rights in the church and no right to receipt of tithes.

⁵⁹ For a discussion of the history of this right see J. C. Fowler, *Church Pews, their Origin and Legal Incidents, with some Observations of the Propriety of Abolishing them, in Three Chapters* (London: Francis and John Rivington, 1844).

⁶⁰ See *Gosling v Veley* (1853) 4 H.L. App. Cas. 679. Though generally the rector was liable for the repair of the chancel (per *Hawkins v Rous* (1695) Holt K.B. 139 cited in W. Dale, *The Law of the Parish Church*, p.72.

numbers of Englishmen, though they did not actively dissent, were distanced from the Church by their own apathy and doubts.⁶¹

The Church experienced the anxieties and difficulties of a church which, though it had begun life as an embodiment of a confessional State, was increasingly alienated from the nation that it aspired to serve. It awoke to the reality that many citizens either rejected its services or, in a hurried and worried existence, did not feel that its services were of immediate use or importance to them. This reality prompted one eminent Churchman to cry:

“What is the national Church? – Who are her members? – Does she number among her members the majority of the people? Let the truth be known. The computation of the numbers said to be within her pale is founded on a fallacy. The number of Papists and other Dissenters having been correctly ascertained, their aggregate is subtracted from the whole population, and the residue are always claimed by the Church of England and Ireland to be her members. It is not possible to conceive a mode of computation more deceptive.”⁶²

The social and political changes of the later nineteenth century, in addition to contributing to a sense of crisis in the Church, also had serious practical consequences for the mechanisms and arrangements by which it had historically supported its national mission. The increasingly non-Anglican character of Parliament led to pressure for mechanisms of Church government and administration independent of it. This culminated in the revival of convocations, diocesan synods, parochial church councils and, ultimately, the formation of the Church Assembly.⁶³

Similar innovation was required in respect of Church maintenance and support. Tithes were devalued by falling agricultural prices and commuted by statute, rendering them of little value. Compulsory church rates were abolished by statute in 1868, and there was a swell of public opinion against the pew rents charged in many parishes, as contributing to the Church's failure to reach the poor.⁶⁴ Further, as the Church sought to respond to the changing social demographics brought about by the Industrial Revolution, ways had to be found of supporting parishes and missions which were not supported by historic endowments. The net result was a greatly increased reliance upon voluntary action and private donations as a means of Church support.

⁶¹ See generally O. Chadwick, *The Victorian Church*, Vol.II (London: A. and C. Black, 1970) and O. Chadwick, *The Secularisation of the European Mind in the nineteenth Century* (Cambridge University Press, 1975).

⁶² R. Oastler, *The Church and the People* (London: C. W. Reynell, 1860) p.13.

⁶³ See e.g. ‘A Few Thoughts on the Laity by a Layman,’ *Contemporary Review*, vol.13 (1870) p.86 and ‘A Speech by J. T. Tomlinson Esq. at the Sheffield Church Congress, 1878,’ in J. T. Tomlinson, *Collected Tracts on Ritual* (vol.I)(London: Church Association, 1899) p.4.

⁶⁴ Though the Church was required by law to reserve free seats to the poor it was felt that the stigma attached to these seats, their poor position in the church, and the perpetuation of privilege all contributed to the alienation of the poor from the Church. See generally O. Chadwick, *The Victorian Church* (Vol.II).

While the means of funding the Church changed so too did the parochial system which they had traditionally supported. One of the Church's first responses to the perceived needs of the rapidly swelling urban population was a Herculean effort to build new churches in the towns and cities. A side effect of this was an increasing mobility among urban congregations. The late nineteenth century was a time of sharp theological divisions within the broad confines of the Church. This, combined with a range of easily accessible churches in any geographical area, served to undermine and dissolve the traditional tie between a parishioner and his parish church. Urban parishioners increasingly chose to attend, not the parish church in respect of which their place of residence gave them rights, but rather the church which best reflected their ecclesiastical tastes and sensibilities.⁶⁵

Another, and related, manifestation, of the changes to the traditional territorial system, was innovation in the mission of the Church to the masses. The urgent need to bring people to salvation, in concert with a profound appreciation of the growing numbers of people living outside of the daily influence of the Church, created this impetus to innovate. Church building schemes alone had not managed to Christianise the masses so other less orthodox strategies were applied. Mission meetings were held at Exeter Hall, services were conducted in parks and theatres, mission churches were established, all in an effort to reach the masses, and all existing outside of the territorial system. As a consequence of this increase in itinerant mission there existed a body of the laity who, though they might have called themselves Churchmen and women, took no part in the parochial, territorial system.⁶⁶

The changes just described distorted and weakened the parochial and territorial system by which the Church had traditionally aspired to fulfil its national mission. More broadly, however, social and political changes raised a profound challenge to the involuntary model of Church membership which underpinned that aspiration. Though the law continued to assert that model of membership,⁶⁷ a sense of alienation from the wider nation, together with initiatives for the formation of representative church councils, and reliance on lay action for Church support and maintenance, contributed to the practice of differentiating between the laity and other citizens. For an increasing number of purposes the Church adopted the voluntary model of membership which typified independent religious bodies and required individuals to take positive steps to assert their membership of the Church. As such, individuals began to experience a differentiation between their duties as lay members of the Church and as citizens of the State.⁶⁸

⁶⁵ For a panoramic picture of the Church at this time see generally O. Chadwick, *The Victorian Church* (Vol.II).

⁶⁶ *ibid.*

⁶⁷ See, e.g. *Escott v Mastin* 4 Moo PCC 104 and *Combe v Edward* (1876) 4 A & E 390.

⁶⁸ See generally M. J. D. Roberts, *The Role of the Laity in the Church of England c1850 – 1885*.

Finally, while many Churchmen strongly defended the national identity and broadly national mission of the Church,⁶⁹ the theological concerns of others placed further pressures on the traditional involuntary model of Church membership. At one extreme of the range of thinking on Church membership were many, but by no means all, members of the Oxford Movement, and those with Ritualist sympathies. For these Churchmen the doctrinal unity and purity of the Church, the spiritual identity of which they emphasised, and for which they sought a source of authority untrammelled by the State, was significantly more important than concerns to maintain a broad and nationally based membership. Their preoccupations with the doctrinal unity of the Church as a pure spiritual body often caused them to take a selective approach to Church membership.⁷⁰

At the other end of the spectrum stood the Evangelicals, whose focus was not the purity or authority of the Church but rather the faith and Salvation of individuals. A central feature of Evangelicalism was the drive to save individual souls. This in itself mitigated against any tendency which might narrow the Church and impair its capacity to reach and serve the masses. In addition Evangelicalism was, as a movement, influenced by forms of theological thought which tempered any disposition to narrow definitions of the Church. Some Evangelicals, for example, were influenced by Calvinistic thought. As such they maintained a clear distinction between the Church on earth - which contained both the elect who would achieve eternal life and the non-elect who had not been saved – and the One Church – which contained only the elect who had been saved. Such Calvinistic doctrine tended in itself to undermine the notion, so essential to the Oxford Movement, of the Church as a body of the faithful. Other Evangelicals, were much influenced by Arminianism, which emphasised the belief that Christ died to save all who believed. This, rather than undermining the vision of the Church as a body of the faithful, made that body as wide and inclusive as possible,⁷¹ thus once again contrasting with the more elitist approach of the Oxford Movement and its successors.⁷²

The Conundrum Created by Establishment and Political and Social Change

The conundrum created by Establishment stemmed from the junction at its heart between Church and State, and the inextricable links that it created

⁶⁹ See, e.g. H. B. Wilson, 'Seances Historiques de Geneve', T Hughes, *This Old Church, What Shall we do with it?* (London: Macmillan, 1878) and a variety of speeches at the Church Congress of 1883.

⁷⁰ See, e.g. *Parl. Deb.* (3rd Ser.) CCXIX, col. 1138 – 39 and *Address Delivered at the Special Meeting of the English Church Union on Tuesday January 16 1877* (London: E.C.U., 1877).

⁷¹ Both Calvinism and Arminianism cut across party lines within the Church and were by no means unique to Evangelicalism.

⁷² Once again constraints on space prohibit a full and detailed account of the Evangelical position and theology however see generally D. W. Bebbington, *Evangelicalism in Modern Britain: A History from the 1730s to the 1980s* (London: Unwin Hymen, 1989), I. Bradley, *The Call to Seriousness: The Evangelical Impact on the Victorians* (London: Cape, 1976), and B. Hilton, *The Age of Atonement: The Influence of Evangelicalism on Social and Economic Thought, 1795-1865* (Oxford: Clarendon Press, 1988).

between Church and State, and Church and nation. This junction created a fundamental ambiguity in the determination of the character and identity of the Church of England. This ambiguity was compounded by the very different reactions of different sections of the Church to the situation in which it found itself.

The Ambiguity Created by the Relationship between Church and State

The Church, as revealed by the survey of the legal aspect of Establishment, was closely tied to the State, and, by virtue of Establishment, had recourse to the civil power in essential matters of its life and discipline. Ecclesiastical laws could, for example, be enforced in the last resort by the coercive powers of the State. Similarly, the final appeal in ecclesiastical causes was to the Crown, while a significant body of ecclesiastical law was enacted by Parliament. In essence Establishment conferred upon the Church two modes or sources of action in respect of its regulation and discipline, and two sources of identity. In many matters it acted through things, such as the pastoral jurisdiction of the bishop, which accrued to it by virtue of its status as a church without reference to the State. In other matters it expressly appealed to things, such as parliamentary legislation and coercive jurisdiction, which were only open to it as a result of its links to the State.

While both Church and State, though in many ways different, shared a spiritual identity, and an essential unity or union as different aspects of the nation, there was little difficulty in the availability of these two distinct courses of action. Though the authority and jurisdiction of Church and State were possessed of different concerns and modes of action, both were, in essence, spiritual. Whether the Church chose to proceed via those powers it had apart from its relationship with the State, or by powers at its disposal by virtue of that relationship, it appealed fundamentally to authority or modes of action which were essentially spiritual or divinely ordained. It was unambiguously a spiritual body.

The ambiguity or conundrum arose when, in the face of constitutional change, some (but not all) parts of the Church lost sight of the shared spiritual identity and unity of Church and State as different aspects of the nation. Church and State remained united in law but this was no longer seen as a union of two types of spiritual institution. It was seen instead as a union of the secular and the spiritual. Moreover, when the Church had recourse to the civil power, for example in legislation or the final appeal to the Crown, this was seen as an appeal to secular power. Thus it appeared that, by virtue of Establishment, there was a junction of the secular and the spiritual, and that the Church had a proper appeal to secular and spiritual authority. Further, it also appeared that the Church, though a spiritual body, was in some respects also a religious (or even secular) institution defined by its relationship with the State.

The Ambiguity Created by the Relationship between Church and Nation

Establishment created a dual identity or personality for the Church, as both a religious institution tied to the State, and as a spiritual body which appealed to God for its authority, wholly independent of the State.⁷³ This duality was mirrored in definitions of the membership of the Church. It could define itself by relation to its geographical relationship with the nation and its role as a national church. Thus *Halsbury's Laws of England*, for example, defined the Church as:

“[T]he branch of the Church which was founded in England when the English were gradually converted to Christianity between the years 597 and 768.”⁷⁴

It could also define itself, not by reference to its territorial limits or national role, but by reference to a core measure of doctrinal unity, or unity in belief, shared by its members. Thus Article XIX of the Articles of Religion defined the Church as:

“A congregation of faithful men, in which the pure word of God is preached, and the Sacraments be duly administered according to Christ's ordinance in all those things that of necessity are requisite to the same.”

While the Confessional State existed, at least in theory, the two types of definition could identify the same persons as members of the Church. This was not the case in the later nineteenth century. As such two models of membership, and two identities, came into conflict. The definition focussed on territorial extent supported an involuntary model of membership, which contributed to the identity of the Church as a national institution. By contrast, a definition focussed on doctrinal unity supported a voluntary model of membership which contributed to the identity of the Church as a private spiritual body or membership society.

Establishment Today: Stasis, Change and a Very Real Conundrum

What follows is a brief description of the constitutional arrangements which today form the Establishment, and a demonstration of the continuing relevance of the problems experienced in the nineteenth century.

Establishment as a Relationship between the Church and the Organs of the State

The doctrine of Royal Supremacy remains at the heart of the modern Church's doctrine of the relationship between Church and State. Canon A7 asserts that the Sovereign:

⁷³ See, for example, S. T. Coleridge, *On the Constitution of Church and State*, chapter five, and R. Palmer, *The Endowment and Establishment of the Church of England*, p.4.

⁷⁴ *Halsbury's Laws of England* (4th ed.) (2000) vol.14, para.345.

“. . . acting according to the laws of the realm, is the highest power under God in this kingdom, and has supreme authority over all persons in all causes, as well ecclesiastical as civil.”

The Royal Supremacy, and the more modern doctrine of Parliamentary Sovereignty, are expressed in all areas of Church administration, government, legislation and adjudication. While there has been an undeniable tendency to distance those matters from State control, and to increase the Church’s abilities of self-determination, there has been no break of the intimate connection between Church and State.

Looking first at the mechanism for appointing Anglican bishops, the essential elements of this system remain unchanged, and the Crown continues to appoint the Church’s bishops. In some respects, however, the practical reality of this system has altered in recent years. Following the inception of the present constitutional arrangements in the Bill of Rights 1688, entailing a constitutional or limited Monarchy, it is the Prime Minister, rather than the Monarch, who appoints bishops. Further, following demands for greater Church influence in the appointment of its bishops, the Prime Minister has, since 1977, been aided by the Crown Appointments Commission. This Church body presents the Prime Minister with two candidates for appointment. The Prime Minister is under no obligation to accept these candidates but, by convention, will normally appoint the first named candidate.⁷⁵ He no longer has a completely free rein in Church appointments.⁷⁶

A similar combination of stasis and change can be seen in relation to Church government and the Church’s powers to make canon law, and to legislate by other means. As such, while the organs of the Church no longer require the licence of the Queen to meet and make canon law, it remains the case, under the Synodical Government Measure 1969 section 1(3)(b), that the General Synod cannot generally pass canons which are repugnant to the royal prerogative or other laws and customs of the realm. This limitation no longer applies to matters of worship or the assent to doctrine.⁷⁷ Ultimately, however, the Church must still obtain the assent of the Queen before its canons can be promulgated.⁷⁸

In respect of Church government, the Church of England Assembly (Powers) Act 1918⁷⁹ created the Church Assembly, a national deliberative and administrative body for the Church outside of Parliament. This trend was continued by the Synodical Government Measure 1969, which established the ‘new and improved’ General Synod.

⁷⁵ See Julian Rivers, ‘From Toleration to Pluralism: Religious Liberty and Religious Establishment under the United Kingdom’s Human Rights Act,’ in R. J. Ahdar (Ed.), *Law and Religion* (Aldershot: Ashgate, 2000).

⁷⁶ It appears though that, however exceptionally, the current Prime Minister has twice breached convention and rejected the Church’s first named candidate. See K. Medhurst, ‘The Church of England: A Progress Report’, (1999) *Parliamentary Affairs* pp.275-290.

⁷⁷ See the Church of England (Worship and Doctrine) Measure 1974 s.6(1).

⁷⁸ See the Synodical Government Measure 1969 s.1(3)(a).

⁷⁹ 9 & 10 Geo. 5 c. 76.

These innovations in the sphere of Church government, and a general dissatisfaction with being reliant upon the interventions of an increasingly busy and pluralistic Parliament, also led to reform in the area of ecclesiastical legislation. The Church of England Assembly (Powers) Act 1918 gave the Church Assembly – and later the General Synod – the power, subject to the final approval of Parliament and the Royal Assent, to pass legislation on matters concerning the Church of England. Under the framework provided by that Act the Legislative Committee of the General Synod lays legislation before the Ecclesiastical Committee of Parliament, which then reports to Parliament on whether the proposed legislation is expedient or not. Parliament can decline to pass the legislation laid before it but it cannot amend the legislation. If the legislation is passed then it becomes a Measure and is a form of primary legislation. Yet, even if the Church has in this way gained a considerable degree of autonomy, it is still ultimately subject to the authority of the Queen in Parliament.⁸⁰

Turning next to the courts one sees perhaps the greatest changes to the traditional embodiments of Establishment. Remnants of the coercive jurisdiction of ecclesiastical courts remain. Under the Ecclesiastical Jurisdiction Measure 1963 s81, for example, an Episcopal Court retains powers to compel the attendance of witnesses and to condemn individuals in costs. However that jurisdiction has very largely passed into history. Similarly, the appellate jurisdiction of the Crown has all but disappeared. Although, under the Ecclesiastical Jurisdiction Measure 1963 section 1(3)(d), the Judicial Committee of the Privy Council remains the court of final ecclesiastical appeal in respect of faculty cases which do not raise matters of doctrine, ritual or ceremonial; jurisdiction (both original and appellate) in matters of worship, doctrine and ritual now rests with the Court of Ecclesiastical Causes Reserved. Though that court is staffed by judges appointed by the Queen, an appeal to it does not constitute an appeal to the Crown.⁸¹

One area of constancy is the status and character of ecclesiastical law. This maintains its heterogeneous character and remains part of the ordinary law of the land. It is not, in the contemplation of the law, the law of a private membership body.

Establishment as a Relationship between Church and Nation

The expression of Establishment as a relationship between Church and nation remains much the same as it did on the late nineteenth century. The Church continues to proclaim a national mission,⁸² and it still seeks to fulfil that mission primarily through the parochial and territorial systems supported by its historic endowments. However, as in the nineteenth century, there is a

⁸⁰ Some scholars have pointed to the reluctance of Parliament to use its veto of Church Measures as an indication that the State will not intervene in Church matters. However others have demonstrated that Parliament, though reluctant to veto, plays an active role in shaping Church legislation. See V. Bogdanor, *The Monarchy and the Constitution* (O.U.P., 1995) ch.9 and J. Behrens, 'The Churchwardens Measure', (2001) 6 Ecc. LJ 97-110.

⁸¹ See the Ecclesiastical Jurisdiction Measure 1963 s.1(3)(b).

⁸² See the Church of England's submission to the Wakeham Commission (Cm 4534, 2000).

recognition that many individuals are alienated from the Church, and that this poses challenges for the Church and its organisation. So, for example, one inner city bishop has recently proclaimed:

“We live on the cusp of modernity and post-modernity. The one will not eradicate the other but our task as a national church is to minister to people for whom there are no metanarratives or paradigms, where the old certainties have disappeared, and for whom the virtue of diversity and choice has become absolute... The task of the Church in worship and mission is to engage with people who are rapidly losing the grammar of the Christian faith (or indeed have never learned it), many of whom are none the less deeply religious.”⁸³

The same bishop, though he clearly stated his commitment to the national mission of the Church and the parochial system by which it was traditionally supported, echoed the frustrations of many of his nineteenth century forebears, stating:

“It is time for the Church of England to reinvent itself and create new paradigms of church while not forsaking the traditional structures. But we cannot do that while the ecclesiastical legal framework is more concerned with the rights of freeholders, patrons and parishioners who never set foot inside their parish church than with the solemn and primary task of making possible the encounter of human beings with the God who created them for a relationship with himself.”⁸⁴

This statement illustrates the modern relevance of the tensions, demonstrated above in relation to the nineteenth century, between the historical position, aspirations and structures of the Church of England, and the current state and needs of the nation which it still aspires to serve.

That tension is further illustrated in current legal approaches to the membership of the Church of England, lay rights and representative Church bodies. The rights of individuals in respect of the services of the Church continue to be defined, *prima facie*, in relation to a geographical relationship with a parish church. Similarly, the law assumes that an individual parishioner who claims rights in respect of a church is a member of the Church of England.⁸⁵ However, recent Measures have recognised the creation of individual rights where, though an individual does not live in a parish, they habitually worship at the church of that parish.⁸⁶ So, too, while the law assumes that an individual is a member of the Church for the purpose of the vindication of basic rights, for other purposes it imposes more stringent requirements upon individuals. Thus, for example, those seeking to

⁸³ The Rt. Revd. Pete Broadbent, Bishop of Willesden, ‘Desperate Measures’, (2002) 6 Ecc. LJ 390.

⁸⁴ *ibid.*, p.394.

⁸⁵ See, e.g. *Marshall v Graham* [1907] 2 KBD 112 at 124 and *Faith v Allen* [1953] Ch.116.

⁸⁶ See, e.g. church Representation Rules, r.1(2)(a) and r.1(2)(b).

be appointed as churchwardens must be baptised, to have their names upon the electoral roll of the parish, and to be actual communicants.⁸⁷

The Modern Relevance of the Conundrum Created by Establishment

The continued relevance of the conundrum created by the combination of Establishment and constitutional and social change in the nineteenth century is self-evident today. Though the State retains many of the trappings of its Christian and confessional past, it is now an avowedly pluralist State. Indeed, there is a commonly expressed belief that it is increasingly secular.⁸⁸ Few today would have sympathy with the claims of the State to be a spiritual body. They might instead be more inclined to favour the sentiment expressed in MacIntyre's assertion that the State is:

“[A] dangerous and unmanageable institution, presenting itself on the one hand as a bureaucratic supplier of goods and services, which is always about to, and never does, give its clients value for money, and on the other as a repository of sacred values, which from time to time invites one to lay down one's life on its behalf. As I have remarked elsewhere . . . it is like being asked to die for the telephone company.”⁸⁹

While the State has lost its spiritual identity, the nation has manifestly become more multicultural and multi-faith than ever before. While the majority of citizens still maintain some links or sense of affiliation with Christianity, the majority of citizens are no longer Anglicans, and a significant minority of them belong to other faith communities.⁹⁰

Establishment remains largely intact. By virtue of it the Church is tied to an increasingly secular State and an increasingly heterogeneous nation. How then is it to define itself? The answer is in a multitude of different ways depending on the context in which it is being required to make that definition, and depending upon the people asked in what is self-consciously a very broad church. The result is an ambiguity which has consequences for the legal and constitutional treatment of the Church of England today.

⁸⁷ Churchwardens Measure 2001 s.1(3). Under the Church Representation Rules R54 defines an actual communicant as a person who has received communion according to the rites of the Church at least three times in the preceding 12 months.

⁸⁸ See, e.g. D. Harte, 'Defining the Legal Boundaries of Orthodoxy for Public and Private Religion in England', in R. O'Dair and A. Lewis (Eds.), *Law and Religion* (Oxford, 2001).

⁸⁹ A. MacIntyre, 'A Partial Response to My Critics,' in J. Horton and S. Mendus (Eds.), *After MacIntyre: Critical Perspectives on the Work of Alistair MacIntyre* (Notre Dame: University of Notre Dame Press, 1994) p.303.

⁹⁰ See the 2001 Census.

A Demonstration of the Conundrum: *Aston Cantlow v Wallbank* and the Treatment of Anglican Institutions and their Functions Under the Human Rights Act section 6

In *Aston Cantlow v Wallbank* the House of Lords, and the Court of Appeal before it, came face to face with the conundrum posed by Establishment in the modern constitution. They were asked to consider the status of Anglican Parochial Church Councils, and the nature of their function in enforcing chancel repair obligations, for the purposes of section 6 of the Human Rights Act 1998. The courts were asked to determine whether PCCs were public authorities, or whether, if not public authorities *per se*, when enforcing an obligation to repair the chancel of a parish church they were performing a function of a public nature. To answer these questions the courts were required to confront the ambiguity surrounding the identity, functions and membership of the Church of England and its institutions.⁹¹

The Majority Decision in the House of Lords

The treatment accorded to the Church of England, chancel repair obligations, and PCCs by the House of Lords was profoundly influenced by European Human Rights Jurisprudence. Following that jurisprudence, and relying upon Article 34 of the ECHR, their Lordships asked the question whether the body concerned was a governmental body, and not, as in the domestic jurisprudence which dominated in the reasoning of the Court of Appeal, whether it was a public body.

The use of the word ‘governmental’ led to a collision with the difficulties and controversies of the late nineteenth and early twentieth centuries. In finding that neither the Church in general, nor PCCs in particular, were governmental bodies their Lordships relied on the famous dictum of Phillimore J that:

“A Church which is established is not thereby made a department of the State. The process of establishment means that the State has accepted the Church as the religious body in its opinion truly teaching the Christian faith, and given to it a certain legal position, and to its decrees, if rendered under certain legal conditions, certain civil sanctions.”⁹²

Their reliance upon this dictum quite reasonably took no account of its history. That history is that it was made by a militant High Churchman, who advocated disestablishment, and protested against all forms of secular interference in the affairs of the Church as an unwarranted intrusion upon its

⁹¹ For a thorough examination of this see A. Pearce, ‘*Aston Cantlow: Chancel repairs and the Status of Church of England Institutions*’, *Law and Justice* 151 (2003). For general critical comment on the decision and HRA 1998 s.6 see F. Meisel, ‘The *Aston Cantlow* Case: Blots on English Jurisprudence and the Public/Private Law Divide’, [2004] P.L. 2-10; L. Turano, ‘Restoring Chancels and Human Rights’, [2004] CLJ 7-9; P. Cane, ‘Church, State and Human Rights: Are Parish Councils Public Authorities’, (2004) 120 LQR 41-48; and D. Oliver, ‘Functions of a Public Nature Under the Human Rights Act’ [2004] PL 329-51.

⁹² *Marshall v Graham* [1907] 2 KBD 112 at 126.

spiritual authority.⁹³ It was made at a time when some politicians and Churchmen still advocated extensive State control of the Church, and the supremacy of secular authority over spiritual concerns.⁹⁴ It is arguable that, though many of these concerns have been mitigated by the reforms of the twentieth century, an unwary reliance upon the dictum led to the House of Lords being unduly reticent in its treatment of Establishment.

A further cause of such reticence might be found more generally in the Human Rights framework within which the court had to determine the case. As many scholars have noted, that framework tends to relegate religion to the private sphere, and to treat it as something with which the State cannot interfere. It does so by setting the individual and the State in opposition to each other, and by framing rights in terms of State duties of non-interference.⁹⁵

The influence of this jurisprudence is clear in the reasoning of the majority of their Lordships. In reasoning from Article 34 they held that neither the Church nor the PCC could be a core public authority because that would mean that, in distinction to other religious bodies, they could not vindicate their Convention rights.⁹⁶ Equally, applying the “governmental test” caused them to assert that the character and functions of the bodies concerned were essentially private, as they were unable to find that the maintenance and provision of religious services was something for which the State would otherwise provide. In sum, it caused the majority, in deferring to the spiritual identity of the Church, to pay scant regard to its public status as a religious institution which is closely tied to the State and exercises a national role and mission.

This tendency to draw a sharp contrast between the secular State and the spiritual and essentially private character and functions of the Church, is reflected in their Lordships’ definitions of Establishment. So, for example, Lord Nicholls maintained a careful distance between Church and State, and a clear distinction between what was governmental and what was spiritual and private, holding:

“Historically the Church of England has discharged an important and influential role in the life of this country. As the established Church it still has special links with central government. But the Church of England remains essentially a religious organisation. This is so even though some of the emanations of the Church discharge functions which may

⁹³ See his book W. G. F. Phillimore, *The Government of the Church in Relation to the State and the Laity* (187? at LPL, reprinted from the *Union Review*).

⁹⁴ See, e.g. the speeches and letters of Sir William Harcourt – a typical example is found in *The Times* 20th July 1874.

⁹⁵ See J. Rivers, ‘From Toleration to Pluralism: Religious Liberty and Religious Establishment under the United Kingdom’s Human Rights Act’, pp.138-62 in R. J. Adhar (Ed.), *Law and Religion* (Aldershot: Ashgate, 2000); S. C. Van Bijsterveld, ‘Religion, International Law and Policy in the Wider European Arena: New Dimensions and Developments’, pp.163-80 in R. J. Adhar (Ed.), *Law and Religion* and S. C. Van Bijsterveld, ‘Freedom of Religion: Legal Perspectives’, pp.299-310 in R. O’Dair and A. Lewis (eds.), *Law and Religion*.

⁹⁶ See [2003] UKHL 37 at para.8.

qualify as governmental . . . This should not be regarded as infecting the Church of England as a whole, or its emanations in general, with the character of a governmental organisation.⁹⁷

Lord Hope, too, carefully preserved the spiritual identity of the Church, holding:

“What establishment in law means is that the state has incorporated its law into the law of the realm as a branch of its general law . . . The Church of England is identified with the state in other ways, the monarch being the head of each . . . It has regulatory functions within its own sphere, but it cannot be said to be part of government. The state has not surrendered or delegated any of its functions or powers to the Church. None of the functions that the Church of England performs would have to be performed in its place by the state if the Church were to abdicate its responsibility . . . The relationship which the state has with the Church of England is one of recognition, not of the devolution to it of any of the powers or functions of government.”⁹⁸

Their Lordships maintained this emphasis upon the Church’s identity as a spiritual body in their treatment of PCCs. As such they focussed on their formulation under Measure and their primary role as being to aid the minister in furthering the mission of the Church in the parish – holding this to be an innately private and spiritual function. In some respects this was well founded, for, as noted, these bodies were the product of a need to provide a forum for Church administration in which only active conformists were able to participate.⁹⁹

Yet their Lordships neglected the intertwining of that “private membership” aspect of PCCs with the historic ideals and framework of a national church. As such they failed to note that the duty to enforce chancel repair obligations came to PCCs from the parish vestries, which exercised civil and spiritual functions and were the very embodiment of the close territorial and administrative relationship between Church, State and nation.¹⁰⁰ This, in turn, caused them to focus upon a characterisation of the Church and PCCs as serving a membership body, rather than the nation as a whole.¹⁰¹

The majority carried this view into their characterisation of the obligation at the heart of the case. The obligation to repair the chancel attached to ownership of rectoral land and was part of the historic system of public taxation by which the State provided for the services of religion in parish churches. It was an obligation undertaken in consequence of the rector’s right to receive tithes. This right was effectively abolished and, following

⁹⁷ *ibid.*, para.13.

⁹⁸ *ibid.*, para.61.

⁹⁹ *ibid.*, para.14, 59, 86 and 156.

¹⁰⁰ This point is noted in the Court of Appeal. See [2001] EWCA Civ 713 at paras. 32-35. See also, A. Pearce, Submission to the Joint Committee on Human Rights – published as an appendix to the 7th report of the 2003-4 session (HL39, HC 382).

¹⁰¹ *Op.Cit.* n.98.

the Chancel Repairs Act 1932, the obligation to repair chancels took the form of private law obligation. Having summarised this history the majority held that the enforcement of the chancel repair obligation was analogous to the enforcement of any other private law obligation or civil law debt. In doing so they dismissed its place in the system of State support which supported the ideal and mission of a national and Established Church.¹⁰²

In short, the treatment accorded to Establishment by the majority in the House of Lords demonstrated a determination to treat the Church of England as it would any other religious body.¹⁰³ Within the Human Rights context this grew out of a focus on its identity as a spiritual body which exists independent of the State and can determine its own actions.

The Dissenting Judgment of Lord Scott

In fact there was one lone voice speaking in the House of the Lords for the traditional identity and role of the Church of England as a religious institution closely linked to the State and ministering to the nation. Lord Scott strongly asserted that traditional role of the Church, the general applicability of rights in respect of its worship and rites as a matter of law, and the history of State interest and support which underpinned them.

Though Lord Scott rejected the argument that PCCs were “core” public authorities for the purposes of HRA s6, he accepted the argument that, in enforcing the obligation to repair the chancel, they were executing a function of a public nature under s6(3)(b). In doing so he asserted strongly the public and national status and mission of the Church under Establishment, holding:

“The important factors and matters relevant to the question in the present case seem to me, in no particular order of importance, to be the following. (1) The parish church is a church of the Church of England, a church by law established. (2) It is a church to which the Anglican public are entitled to have recourse, regardless of whether they are practising members of the church, for marriage, for baptism of their children, for weddings, for funerals and burial, and perhaps for other purposes as well. (3) Members of other denominations, or even of other religions, are, if parishioners, entitled to burial in the parish churchyard. (4) The church is, therefore, a public building. It is not a private building from which the public can lawfully be excluded at the whim of the owner. (5) . . . (6) A decision by a PCC to enforce a chancel repairing liability is a decision taken in the interests of the parishioners as a whole. It is not taken in pursuit of any private interests. If it were so taken, it would I think be impeachable by judicial review.”¹⁰⁴

This characterisation of the Church caused him to reject the notion that the enforcement of the chancel repair obligation was a private matter, because

¹⁰² [2003] UKHL 37 at paras.16, 64, and 89.

¹⁰³ Indeed the majority took as authority for the private status of religious bodies the *Wauchmann Case* ([1993] 2 All ER 249) which concerned a Jewish body which enjoyed no formal constitutional relationship with the State.

¹⁰⁴ [2003] UKHL 37 at para.130.

enforcement was undertaken on behalf of all ‘parishioners’ and not on behalf of a congregation or private membership body.

The Court of Appeal

Like Lord Scott the Court of appeal had emphasised the public and national character and status of the Church of England and its institutions and their functions. At every stage of the judgment of the court Sir Andrew Morritt highlighted the dual character of the Church and its institutions in consequence of Establishment. Thus, for example, he held:

“The Church of England has enjoyed a unique status . . . since the passage in 1532-4 of the five statutes which severed the hegemony of Rome and placed the Church under the spiritual and temporal sovereignty of the Crown; this notwithstanding its theological continuity since Saxon times . . . Both the spiritual and the temporal courts were thenceforth the King’s courts, and it is by the latter that the liability of first spiritual and then lay rectors to repair the chancel has been both enunciated and enforced.”¹⁰⁵

Similarly, he recognised that PCCs had both civil and spiritual functions, asserting:

“The PCC itself exemplifies the special status of the Church of which it forms part . . . The successor of the vestry, the PCC is constituted not as a voluntary association but by law. In the exercise of statutory powers originating . . . in the Act of Supremacy 1568, but found more immediately in the Church of England Assembly (Powers) Act 1919, the Parochial Church Councils (Powers) Measure 1956 as amended provides for the discharge of the functions of the PCC and, by s3, constitutes it a body corporate. That a measure of its National Assembly can make every PCC a statutory corporation is an index both of the public character of the Church of England and (to some extent) of the PCC. The functions of the latter, which include but are not confined to those set out in s2 of the Measure, subsume those of the vestry (s4) which included both spiritual and civil matters.”¹⁰⁶

This approach and emphasis differed greatly from that of the House of Lords. One possible explanation for this is that the Court of Appeal focussed largely on domestic jurisprudence on the public-private divide.¹⁰⁷ As such they were not confronted with the application of words such as “governmental” in respect of the Church or its institutions. Further, it is arguable that they were set at a greater remove from the discourse which today tends to allocate religion to the private sphere.

¹⁰⁵ [2001] EWCA Civ 713 at para.31.

¹⁰⁶ *ibid.*, para.32.

¹⁰⁷ See e.g. *ibid.*, paras.34 and 35.

Conclusion

This article has sought to examine the social, political, theological and constitutional and historical forces which continue today to contribute to a profound difficulty in legal and constitutional attempts to address the position and treatment of the Church of England.

It has been contended that Establishment, by creating a junction between Church, State and nation, has created a source of ambiguity in allocating to the Church diverse identities and appearances. The conundrum which it presents is exacerbated both by the apparently secular or plural identity of the State and nation to which the Church is joined, and the diversity of reactions to that junction. As the reasoning of the House of Lords in *Aston Cantlow v Wallbank* demonstrates, old difficulties are rendered more acute as a result of the Human Rights context within which the courts now operate. This brings into play distinctions and attitudes which are not contemplated within the historic character and aims of Establishment. The result is a fundamental inconsistency in our current constitutional arrangements. The House of Lords showed an unwillingness to recognise and accord significance to the national and public role of the Church. Yet in our arrangements for the coronation of our Monarchs, the composition of the second chamber, our blasphemy laws, Episcopal appointments, and national religious events we continue – for the present at least – to recognise that very public and national character.

RESOLVING THE DURESS DILEMMA: GUIDANCE FROM THE HOUSE OF LORDS

*Ray Ryan, Barrister-at-Law and Des Ryan, Balliol College,
Oxford**

“. . . [Y]our Lordships should hesitate long lest you may be inscribing a charter for terrorists, gang-leaders and kidnappers.” – *Per* Lord Simon of Glaisdale in *Director of Public Prosecutions for Northern Ireland v Lynch* [1975] AC 653, 688.

Introduction

The defence of duress has long been regarded as a notoriously difficult area of the criminal law. Public policy considerations have demanded that for the defence to be successfully pleaded, a high threshold must be crossed. Beyond this, however, there is a paucity of clear guidance from the courts as to the situations in which a defence of duress will be successfully invoked by accused persons and there is thus a need for detailed consideration of the scope and rationale of the defence. With its decision in March 2005 in the case of *R. v Hasan*,¹ the House of Lords has done much to bring greater clarity to the law on duress. The case is significant, *inter alia*, because the House expressly disapproved a decision of long standing with which criminal lawyers will be familiar, the Court of Appeal decision in *R v Hudson and Taylor*.² Moreover, the House of Lords also considered the cases of *Attorney General v Whelan*³ and *R. v Fitzpatrick*,⁴ decisions of the former Irish Court of Criminal Appeal and the Northern Ireland Court of Criminal Appeal respectively. The authoritative formulation of the defence enunciated by Murnaghan J in *Whelan* has enjoyed a remarkable degree of endorsement across the common law world.⁵ In the recent *Hasan* case,

* The authors wish to thank Dr. John E. Stannard for his many very helpful suggestions and comments on an earlier draft. Any remaining errors are the sole responsibility of the authors.

¹ [2005] UKHL 22; [2005] 2 WLR 709 (hereafter '*Hasan*'). It should be noted that the appeal to the House of Lords in *Hasan* also involved an important question as to the meaning of a confession under the Police and Criminal Evidence Act 1984, addressed in the opinion of Lord Steyn, with which all of their Lordships agreed. For present purposes, however, our analysis is confined to the speeches of Lord Bingham of Cornhill and Baroness Hale of Richmond on the issue of duress. Lords Steyn, Rodger of Earlsferry and Brown of Eaton-under-Heywood agreed with the opinion of Lord Bingham.

² [1971] 2 QB 202.

³ [1934] IR 518.

⁴ [1977] NI 20.

⁵ Examples include the approval of *Whelan* in the speeches of members of the Appellate Committee of the House of Lords in *Director of Public Prosecutions for Northern Ireland v Lynch* [1975] AC 653. Several Australian texts also refer to the judgment delivered by Murnaghan J, with Gillies regarding it as a "classic statement" (Gillies, *Criminal Law* (4th ed., 1999, p.341); the case is also discussed,

however, one aspect of the approach in *Whelan* has been disapproved. The House of Lords has firmly rejected the view, apparently espoused⁶ in *Whelan*, that the defence of duress can amount to a justification for the conduct of the accused; rather the defence is to be viewed as merely excusing the criminal conduct. Conversely, the *Fitzpatrick* decision has been roundly endorsed by the Appellate Committee in *Hasan*. More generally, it will be seen that the approach of the House of Lords in *Hasan* represents a policy-driven response, aimed at confining this most controversial of defences within a restricted ambit. The purpose of the present article is to record the important observations of the Appellate Committee in this case, and to reflect on the potential influence that the decision may exert upon the operation of the defence of duress in criminal law.

***Hasan*: The Background**

The facts of the *Hasan* case may be stated within a narrow compass. The defendant was charged with aggravated burglary, and relied on the defence of duress. He claimed to have been coerced into committing the burglary by one S, a character known for his violent disposition and his involvement in dealing drugs. At his trial, the judge directed the jury that the accused could not rely on the defence of duress if, in their view, he had, by freely associating with S, run the risk of being subjected to threats. The jury convicted the defendant. On his appeal to the Court of Appeal,⁷ the defendant's conviction was quashed, being regarded as unsafe on the basis that, *inter alia*, the judge had misdirected the jury on the issue of duress. Specifically, the Court of Appeal found fault with the refusal by the trial judge to direct the jury that the accused could rely on duress if they were satisfied that he had not anticipated that he would be compelled to commit a crime as serious as that of aggravated burglary. This approach was perhaps overly generous to accused persons who had chosen to consort with criminals and then sought to excuse their own criminality by stating that those criminals had coerced them into wrongdoing. Public policy, it might be objected, requires that the test for duress impose a more stringent standard on persons having recourse to this defence. It may not have come as much of a surprise then, that on the prosecutor's appeal, a unanimous House of Lords reversed the Court of Appeal, and restored the defendant's conviction.

Judgment of the House of Lords

In the House of Lords, the leading opinion was delivered by Lord Bingham, who conducted a detailed analysis of the history and rationale for the defence of duress as it has evolved and come to be understood over several centuries. Lord Bingham noted at the outset of his opinion the salient distinguishing features which set the defence apart from other defences available in the

albeit more critically, by Bronitt and McSherry, *Principles of Criminal Law* (2001), p.310.

⁶ The ambiguity surrounding this aspect of the *Whelan* formulation can be traced to the penultimate paragraph of the opinion of Murnaghan J in that case, who at p.526 of the report employs, apparently interchangeably, the terms 'justification' and 'excuse'.

⁷ Rix LJ, Crane J and Judge Maddison: [2003] EWCA Crim 191; [2003] 1 WLR 1489 (sub nom *R v Z*).

criminal law. These include the fact that it is now regarded as an excusatory, rather than a justificatory defence, a proposition for which his Lordship cited, *inter alia*, the decision of the House of Lords in *Director of Public Prosecutions for Northern Ireland v Lynch*.⁸ On this point, Lord Bingham considered the judgment of the former Irish Court of Criminal Appeal in *Attorney General v Whelan*,⁹ the classic Irish authority on the defence of the duress. In an oft-cited passage in his judgment in the *Whelan* case, Murnaghan J provided the following definition of the defence:

“It seems to us that threats of immediate death or serious personal violence so great as to overbear the ordinary power of human resistance should be accepted as a justification for acts which would otherwise be criminal.”¹⁰

In *Hasan*, Lord Bingham viewed this formulation as representing a now outmoded approach to the defence, since the better view now appears to be that the defence affords an excuse rather than a justification for unlawful conduct.¹¹ Other distinguishing features of the defence which Lord Bingham emphasised included the fact that when successfully pleaded it acts as a complete, rather than a reductive, defence, totally exonerating the accused; that the victim may ordinarily be assumed to have been morally blameless; and that where the evidence is sufficient to raise the question of duress, the burden rests with the prosecution to establish to the criminal standard that the defendant did not act under duress.¹² In light of these distinguishing features of the defence, Lord Bingham confessed that he found it “unsurprising that the law in this and other jurisdictions should have been developed so as to confine the defence of duress within narrowly defined limits”,¹³ and even went so far as to observe that these very features of the defence inclined him

⁸ [1975] AC 653.

⁹ [1934] IR 518.

¹⁰ *ibid.*, at 526.

¹¹ This criticism of the *Whelan* formulation had been hinted at by Lords Morris of Borth-y-Gest and Edmund-Davies in the course of their opinions in the *Lynch* case: [1975] AC 653, at 673 and 711 respectively. For valuable comment on the conceptual confusion that has prevailed in the case law see Smith, “Duress and Steadfastness: In Pursuit of the Unintelligible” [1999] *Crim LR* 363, 369. On the debate between adopting a justificatory or excusatory rationale in the context of the defence of duress generally see McAuley and McCutcheon, *Criminal Law: A Grammar* (2000), pp.786-817. In the course of her opinion in *Hasan* Baroness Hale noted that the “moral basis of the defence remains a hot topic” ([2005] UKHL 22, [73]), citing the recent article by Wilson, “The Structure of Criminal Defences” [2005] *Crim LR* 108.

¹² Citing *Director of Public Prosecutions for Northern Ireland v Lynch* [1975] AC 653. As is well known, the Law Commission in its Report, *Legislating the Criminal Code: Offences against the Person and General Principles* (1993, Law Com. No 218, Cm 2370, paras.33-34), recommended that a legal burden of proof, on the balance of probabilities, be placed on a defendant to establish a defence of duress. Lord Bingham in his opinion in *Hasan* appeared to regard (at [20]) such a shift as a matter for the legislature and thus beyond the purview of judicial intervention, a point accepted by Baroness Hale – herself one of the signatories to the 1993 Report – at [74]. However, it is clear that the present position, imposing on the Prosecution the burden of disproving duress, comprises another reason for confining the operation of the defence within tight boundaries.

¹³ [2005] UKHL 22, [21].

“where policy choices are to be made, towards tightening rather than relaxing the conditions to be met before duress may be successfully relied on”.¹⁴ His Lordship proceeded to enumerate no less than seven such limitations which have been placed on the ambit of the defence.¹⁵ The list is worth sketching out briefly:

- 1) Duress does not afford a defence to a charge of murder¹⁶ or attempted murder;¹⁷
- 2) The threat forming the basis of the duress must be one to cause death or serious personal injury;
- 3) The threat must be directed against the accused or his immediate family or a person close to him;
- 4) The test for establishing duress –unlike many other criminal law defences– is typically stated in objective terms;
- 5) The defence will only be available where the criminal conduct with which the accused is charged has been directly caused by the threats upon which he relies;
- 6) The accused must not have been in a position where a course of “evasive action” was open to him to take;
- 7) The accused cannot invoke the defence of duress where he has voluntarily laid himself open to that duress.

It is submitted that, of the above limitations on the defence, it is limitations 4, 6 and 7 which are perhaps the most problematic, and it is therefore understandable that these elements received greatest treatment in the opinion of Lord Bingham. It is useful to analyse each in turn.

Objectivity

In considering the nature of the “objectivity” standard by which the criminal law adjudicates upon a defendant’s plea of duress, two inquiries must be distinguished. The first concerns the determination of whether the accused was indeed motivated to act as he did out of fear that the threatened harm would be carried out if he failed to comply. On this point, Lord Bingham in *Hasan* approved the judgment of the Court of Appeal in *R. v Graham*,¹⁸ which had adopted a formulation of “reasonable belief”.¹⁹ The objective element of the test adopted in *Graham* appears to have been somewhat diluted by the subsequent Court of Appeal decision in *R. v Martin*²⁰, where Mantell LJ opined²¹ that, consistent with the *Graham* approach, the question

¹⁴ *ibid.*, at [22].

¹⁵ *ibid.*, at [21].

¹⁶ *R. v Howe* [1987] AC 417, noted Gearty, “Howe To Be A Hero” [1987] *CLJ* 203; Smith, “Must Heroes Behave Heroically?” [1989] *Crim LR* 622.

¹⁷ *R. v Gotts* [1992] 2 AC 412.

¹⁸ [1982] 1 WLR 294, approved by the House of Lords in *R. v Howe* [1987] AC 417.

¹⁹ [1982] 1 WLR 294, 300, *per* Lord Lane CJ.

²⁰ [2000] 2 Cr App Rep 42.

²¹ *ibid.*, at 49.

of the defendant's reasonable belief could be judged in subjective terms.²² Perhaps surprisingly, *Martin* was not considered by Lord Bingham in his opinion in *Hasan*; it nevertheless seems correct to conclude that Mantell LJ's approach in that case cannot be reconciled with that of the Appellate Committee in *Hasan*.

Lord Bingham noted that the relevant Judicial Studies Board Direction,²³ upon which the trial judge in *Hasan* had relied, contained the words "he genuinely believed", but made no reference to reasonableness. Lord Bingham explained that whilst it "is of course essential that the defendant should genuinely, i.e. actually, believe in the efficacy of the threat by which he claims to have been compelled",²⁴ this does not provide a "warrant for relaxing the requirement that the belief must be reasonable as well as genuine".²⁵ It is perhaps salutary to recall that the distinction between a *genuine* and a *reasonable* belief is one which has proved not unproblematic in other areas of the criminal law, notably resurfacing of late in a series of Court of Appeal judgments concerning the drawing of adverse inferences from the silence of the accused under section 34 of the Criminal Justice and Public Order Act 1994;²⁶ the continued use of these overlapping, if not altogether tautologous, terms in the sphere of the objective test in duress, is to be regretted.

The second inquiry which, prior to *Hasan*, had remained unaddressed, was whether the defendant's foresight should be judged by an objective or subjective test, that is, does the defendant lose the benefit of a defence of duress only if he actually foresaw the risk of coercion or does he lose it if he ought reasonably to have foreseen the risk of coercion, whether he actually foresaw the risk or not? Whilst acknowledging that the application of an objective test to other elements of the defence of duress had attracted criticism,²⁷ Lord Bingham nevertheless concluded that policy concerns pointed towards preferring an objective test. This is doubtless due to the distinguishing features of duress outlined by Lord Bingham at the outset of his opinion, features which in his Lordship's view rendered unhelpful analogies based on other defences such as self-defence and provocation. His Lordship thus answered the relevant certified question as follows:

"[T]he defence of duress is excluded when as a result of the accused's voluntary association with others engaged in criminal activity he foresaw or ought reasonably to have

²² His Lordship drew support for this position from the fact that Lord Lane CJ in *Graham* had drawn an analogy between duress by threats and self defence; he also relied on the then very recent decision of the Court of Appeal in *R. v Cairns* [1999] 2 Cr App Rep 137.

²³ Direction No. 49, issued August 2000, available at <http://www.jsboard.co.uk>.

²⁴ [2005] UKHL 22, [23].

²⁵ *ibid.*

²⁶ *R. v Betts and Hall* [2001] EWCA Crim 224, [2001] 2 Cr App Rep 16; *R. v Howell* [2003] EWCA Crim 1, [2005] 1 Cr App Rep 1; *R. v Hoare and Pierce* [2004] EWCA Crim 784, [2005] 1 WLR 1804; *R. v Beckles* [2004] EWCA Crim 2766, [2005] 1 WLR 2829. For incisive comment see Barsby and Fitzpatrick [2005] *Crim LR* 560.

²⁷ [2005] UKHL 22, [38].

foreseen the risk of being subjected to any compulsion by threats of violence.”²⁸

This confirmation of the application of an objective test to these elements of the defence of duress is arguably one of the most important aspects of the *Hasan* decision. Landmark shifts towards subjectivity in the criminal sphere in recent years – the clearest example is perhaps *R. v G*²⁹ – might perhaps have given cause to conjecture that in the context of the defence of duress a subjective approach would ultimately be favoured. With its decision in *Hasan*, the House has robustly quelled any such speculation. The preservation of the objective standard is indicative of the hard line being taken towards the defence and rebuts any concern that the criminal law would resort to what has been termed a “sell-out to subjectivism”.³⁰

The accused must not have been in a position where a course of “evasive action” was open to him to take

Lord Bingham noted that the recent trend in the English authorities had been to emphasise the requirement that the accused should not have been capable of adopting a course of evasive action, his Lordship citing a well-known passage from the opinion of Lord Morris of Borth-y-Gest in *Lynch*.³¹ Lord Morris had cautioned that the defence of duress “must never be allowed to be the easy answer of those who can devise no other explanation of their conduct nor of those who readily could have avoided the dominance of threats nor of those who allow themselves to be at the disposal and under the sway of some gangster-tyrant”.³² However, Lord Bingham expressed his concern that the requirement was one which had been “unduly weakened”³³ of late, and one of the most significant features of the *Hasan* ruling is its express repudiation of any such a tendency to dilute the strictness of this requirement.

The decision of the Court of Appeal in *R. v Hudson and Taylor*³⁴ is one which will be familiar to criminal lawyers. The decision is popularly regarded as having embraced a generous view of the ambit of the duress defence: indeed, Lord Bingham in his opinion in *Hasan* notes Glanville Williams’s description of *Hudson and Taylor* as “an indulgent decision”.³⁵ The appellants were two teenage girls who had committed perjury at an earlier trial by failing to identify the defendant. In the course of their trial for

²⁸ *ibid.*, at [39].

²⁹ *R. v G* [2003] UKHL 50; [2004] 1 AC 1034. In this judgment of a unanimous House of Lords, in which the leading opinion was again delivered by Lord Bingham, the notorious decision of the House in *R v Caldwell* [1982] AC 341 was overruled. See Ryan and Ryan, “Recklessness, Subjectivity and the Criminal Law” (2004) 22 *ILT* (n.s.) 90; but see Kimel (2004) 120 *LQR* 548.

³⁰ This phrase was employed by Horder in his analysis of the Law Commission’s Report, above n.12; see Horder, “Occupying the Moral High Ground? The Law Commission on Duress” [1994] *Crim LR* 334, at 340.

³¹ [1975] AC 653.

³² *ibid.*, at 670.

³³ [2005] UKHL 22, [21].

³⁴ [1971] 2 QB 202.

³⁵ [2005] UKHL 22, [27] citing Williams, *Textbook on Criminal Law* (2nd ed., 1983), p.636.

perjury they pleaded duress on the basis that they had been warned by a group, including a man with a reputation for violence, that if they identified the defendant they would be killed. The trial judge ruled that the threats lacked sufficient immediacy to sustain the defence of duress: the girls could have sought the protection of the police, thereby “neutralising” the threat. This decision was reversed by the Court of Appeal, since although the threats could not be executed in the courtroom they could be carried out in the street that very same night. The Crown’s argument that the appellants had failed to neutralise the threat by seeking police protection was rejected as failing to distinguish between cases in which the police would be in a position to provide effective protection and cases in which they would not.

Lord Bingham expressed his dissatisfaction with *Hudson and Taylor* in very clear terms:

“I can understand that the Court of Appeal. . . had sympathy with the predicament of the young appellants but I cannot, consistently with principle, accept that a witness testifying in the Crown Court at Manchester has no opportunity to avoid complying with a threat incapable of execution then and there.”³⁶

His Lordship further described the decision as having “had the unfortunate effect of weakening the requirement that execution of a threat must be reasonably believed to be imminent and immediate if it is to support a plea of duress”.³⁷ Thus, although not expressly overruled in *Hasan, Hudson and Taylor* was the subject of such disapproving comment as to effectively render the decision no more than a historical anomaly.

The accused cannot invoke the defence of duress where he has voluntarily laid himself open to that duress

It is well established that an accused person cannot invoke the defence of duress where he has voluntarily exposed himself to the threat of which he now complains. Indeed, in the course of her speech in *Hasan*, Baroness Hale said of this question that “[l]ogically, if it applies, it comes before all the other questions raised by the defence”.³⁸ The best-known authority in this regard is perhaps *R. v Fitzpatrick*.³⁹ The accused was charged with murder, robbery and membership of a proscribed organisation. He was a member of the Official IRA, but testified that he had attempted to leave but had been prevented from doing so by threats of violence to himself and to his parents. At trial, the judge ruled that the defence of duress could not apply since the accused had voluntarily exposed himself to the threat. The Northern Ireland Court of Criminal Appeal upheld this decision, Lord Lowry LCJ explaining:

“This court is satisfied that there are circumstances in which persons who associate with violent criminals and voluntarily expose themselves to the risk of compulsion to commit criminal acts cannot according to the common law avail

³⁶ [2005] UKHL 22, [27].

³⁷ *ibid.*

³⁸ *ibid.*, at [74].

³⁹ [1977] NI 20.

themselves of the defence of duress...A person may become associated with a sinister group of men with criminal objectives and coercive methods of ensuring that their lawless enterprises are carried out and thereby voluntarily expose himself to illegal compulsion, whether or not the group is or becomes a proscribed organisation.”⁴⁰

The approach in *Fitzpatrick* was arguably supported by dicta of members of the House of Lords in *Lynch*⁴¹ and was considered by Hutton J (as he then was) in the later case of *R. v Calderwood*.⁴² Calderwood and Moore, as members of a proscribed paramilitary group, had committed a murder. They sought to excuse their behaviour by reliance upon the defence of duress. It was held that that defence was not open to them,⁴³ because by becoming members of an organisation whose objectives were criminal and coercive in nature, the defendants had chosen to expose themselves to the risk of illegal compulsion. *Fitzpatrick* was adopted by the English Court of Appeal in *R. v Sharp*,⁴⁴ a decision cited with approval by Lord Bingham in *Hasan*. In *Sharp*, the appellant was a member of a gang which had carried out a series of armed robberies. He sought to plead duress as a defence to manslaughter when a sub-postmaster was shot dead by the gang ringleader in the course of a robbery. Sharp alleged that he had attempted to withdraw from this robbery when he realised that guns would be used, but that a gun had then been pointed at him and a threat made to “blow his head off” if he did not participate. Giving the judgment of the Court of Appeal, Lord Lane LCJ said:

“... [W]here a person has voluntarily, and with knowledge of its nature, joined a criminal organisation or gang which he knew might bring pressure on him to commit an offence and was an active member when he was put under such pressure, he cannot avail himself of the defence of duress.”⁴⁵

⁴⁰ *ibid.*, at 33.

⁴¹ [1975] AC 653. See for example the opinion of Lord Edmund-Davies at 705.

⁴² [1983] 10 NIJB.

⁴³ *Calderwood*, of course, was decided at a time when there was some confusion as to whether the defence was considered to be available on a charge of murder. According to the House of Lords in *Lynch*, duress was available on the charge of murder to an accomplice or principal in the second degree, but in the course of its advice in *Abbott v The Queen* [1977] AC 755 the Judicial Committee of the Privy Council denied, by a majority, the defence to the principal in the first degree. However, while the two dissenting judges in *Abbott*, Lords Wilberforce and Edmund-Davies, had both been members of the majority in *Lynch*, of the judges in the majority in *Abbott*, only Lord Kilbrandon had formed part of the panel in *Lynch*, and his Lordship had in that case issued a terse dissent. Moreover, as Wilson shrewdly points out, the two minority judges in *Abbott* were “the only specialist criminal lawyers in the court” (Wilson, *Criminal Law Doctrine and Theory*, 2nd ed., 2003, p.276). Thus *Calderwood* was decided against the backdrop of considerable uncertainty as to the availability of the defence of duress to an accused charged with murder. The point was not finally resolved till the unanimous decision of the House in *Howe*: see above n.16.

⁴⁴ [1987] QB 853.

⁴⁵ *ibid.*, at 861. It is apparent from the above passage that the criminal gang or organisation must be one likely to exercise duress and the accused must be aware

However, the controversial question surrounding this limitation on the defence is whether it is sufficient for the prosecution to show merely that the accused knew that he might be compelled to participate in any form of criminal activity, or whether it must be shown that the accused was aware that he would be forced into committing a *particular type* of offence. The Court of Appeal in *Hasan* had taken the view that, in order for this restriction to operate so as to defeat the defence being available, the risk of compulsion of which the accused was aware must be related to “offences of the type charged”. The underlying question for the jury to ask was thus, according to Rix LJ, “whether the accused anticipated that association with criminals could link him, even against his will, with criminality of the seriousness with which he is charged”.⁴⁶ On appeal, the House of Lords rejected this limitation. Lord Bingham noted that there had been “no warrant” for such a gloss on the voluntary association limitation until the judgment of the Court of Appeal in *R. v Baker and Ward*,⁴⁷ a judgment which his Lordship found not to have received widespread application or support in subsequent Court of Appeal decisions,⁴⁸ and one which his Lordship ultimately concluded had “misstated the law”.⁴⁹ His Lordship expressed his rejection of the “type of offence” gloss on the voluntary association limitation in the following robust language:

“The defendant is seeking to be wholly exonerated from the consequences of a crime deliberately committed. The prosecution must negative his defence of duress, if raised by the evidence, beyond reasonable doubt. The defendant is, *ex hypothesi*, a person who has voluntarily surrendered his will to the domination of another. Nothing should turn on foresight of the manner in which, in the event, the dominant party chooses to exploit the defendant’s subservience.”⁵⁰

Baroness Hale in her judgment, while agreeing with Lord Bingham, added an important (if obvious) qualification to the strictness of this approach. Baroness Hale stressed the importance of establishing that the accused had indeed set up “a voluntary association with others”, stating that she did not believe this limitation on the defence to be aimed at such defendants as battered wives or those in close personal and family relationships with their duressors.⁵¹ For Baroness Hale, the other restrictions which operate on the defence of duress are “more than adequate to keep it within bounds in such cases”.⁵²

of this when he joins, as demonstrated by the Court of Appeal decision in *R. v Shepherd* (1988) 86 Cr App Rep 47.

⁴⁶ [2003] EWCA Crim 191, [2003] 1 WLR 1489, [73].

⁴⁷ [1999] 2 Cr App Rep 339.

⁴⁸ *R. v Heath* [2000] *Crim LR* 109; *R. v Harmer* [2002] *Crim LR* 401.

⁴⁹ [2005] UKHL 22, [37].

⁵⁰ *ibid.*

⁵¹ *ibid.*, at [78].

⁵² *ibid.*

Conclusion

The defence of duress has rightly been described as “an extremely vague and elusive juristic concept”.⁵³ With its decision in the *Hasan* case, the House of Lords has provided a clear restatement of the law relating to the operation of the defence. It seems fair to describe the approach adopted by the Appellate Committee as one of extreme caution, predicated upon a fundamental unease with the operation of the defence and a consciousness of the need to guard against its potential abuse by an unworthy defendant. It is significant that the Appellate Committee regards the defence as entirely separate and distinct from the panoply of other defences in the criminal law: as Lord Bingham states, “[t]he only criminal defences which have any close affinity with duress are necessity, where the force or compulsion is exerted not by human threats but by extraneous circumstances, and, perhaps, marital coercion”.⁵⁴

The variety of reasons justifying such a restrictive approach to the ambit of the defence, as articulated by Lord Bingham in his opinion, undoubtedly give rise to concerns that the defence should not be developed beyond its strict parameters. Of particular note here is the apparent support of the Appellate Committee in *Hasan* for the notion that the strictness of these parameters can be supported when one bears in mind the fact that judges can have regard to the question of duress at the sentencing stage.⁵⁵ Thus it appears that the focus towards the conceptual underpinnings of the defence is shifting back, at least in part, to the approach advocated by Lord Simon in his powerful dissenting judgment in *Lynch*, where his Lordship expressed the view that duress should go to mitigation, but not liability.⁵⁶

The strictness of the course adopted in *Hasan* – most notably with regard to the rejection of a subjective test and the staunch adherence of the House to the *Fitzpatrick* doctrine – will doubtless be viewed by some as an overly-harsh response to the legitimate concerns tabulated by Lord Bingham. While it is undoubtedly the case that “the policy of the law must be to discourage association with known criminals, and it should be slow to excuse the criminal conduct of those who do so”,⁵⁷ will the approach adopted to the defence of duress generally in *Hasan* prove to have adequately resolved the duress dilemma?

⁵³ *Per* Lord Simon of Glaisdale in the *Lynch* case: [1975] AC 653, 686.

⁵⁴ [2005] UKHL 22, [19]. As to the impact of *Hasan* on the defence of necessity, it should be noted that Mance LJ (as he then was), speaking for a unanimous Court of Appeal in *Quayle and ors v R*. [2005] EWCA Crim 1415 was guided by the opinion of Lord Bingham in *Hasan* in adopting a similarly strict approach to the appellants’ pleas of necessity of circumstances.

⁵⁵ See especially the opinion of Lord Bingham at para.[22].

⁵⁶ For comment see Glazebrook, “Committing Murder under Duress” [1975] *CLJ* 185.

⁵⁷ *Per* Lord Bingham in *Hasan* [2005] UKHL 22, at [38].

BOOK REVIEW

DEVOLUTION, LAW MAKING AND THE CONSTITUTION.
Edited by Robert Hazell and Richard Rawlings (Imprint Academic,
2005. Hardback, 338 pages, £35.00)

The devolution project in the United Kingdom, which was introduced as part of a wider programme of constitutional change following the election of New Labour in 1997, was not an entirely novel idea as it was predated by periods of devolved government in Northern Ireland, and earlier attempts to devolve power to Scotland and Wales. Nonetheless, its present incarnation can be seen as a fundamental change in the way power is distributed within the United Kingdom. Although the traditional notion of parliamentary sovereignty has not been legally restricted by the establishment of devolved institutions, politically, the freedom of Westminster to legislate on matters which have been devolved, without the consent of the devolved region, is now more constrained. The devolution programme in the United Kingdom is asymmetrical, with each of the devolved regions being granted power to legislate in different areas and with legislation of differing status. Therefore, the impact of devolution on law making within the UK is not uniform across all the nation's constituent parts and cannot be subject to generalisation.

In "Devolution, Law Making and the Constitution", the authors attempt to tackle the question of how far devolution has altered the law making process within the United Kingdom to date by assessing the legislative powers of each of the institutions and the powers retained by central UK authorities. The book opens with an introductory note that provides a useful and clearly explained overview of the differing powers and structures of the parliaments and assemblies of the United Kingdom, and the mechanisms in which disputes between these bodies can be avoided or resolved. The subsequent chapters in the first half of the book focus on the constitutional development in the three devolved regions. In the latter half of the book the focus shifts to the experience within the central UK institutions, with chapters on Whitehall, Westminster and the courts, before the final chapter in which Hazell discusses the need for closer links between the executives and assemblies of the devolved regions and the United Kingdom in order to make devolution more effective. Throughout each of the chapters in this book, it is possible to identify some overarching themes that will be discussed below, and some general conclusions that have been made by the authors. These include a belief that devolution in its current form does not represent a radical shake up of how law is made in the United Kingdom as Westminster has retained supremacy and the ability to legislate on devolved issues; that English issues still dominate in UK-wide legislation; and that the methods of formulating and enacting legislation in the central UK institutions have yet to fully adapt to the changes brought by devolution. Nonetheless, the authors do highlight progress towards the development of distinctive legal cultures in the

devolved regions, and make valuable suggestions on how this process could be continued.

This book resulted from a four-year research project entitled “Law and Devolution Disputes” funded by the ESRC which aimed to assess the “role of the law in shaping the devolution settlement” during its first five years. The issue was analysed “in terms of adjustments to the division of powers between the UK and devolved governments, through legislation, legal interpretation by governments, intergovernmental negotiation and court decisions”. The project entailed a collaboration between the Constitution Unit of UCL and six constitutional and public law experts in Scotland, Wales and Northern Ireland. The empirical data for the project came from interviews with government lawyers in the devolved and the central UK institutions. The resulting text provides a thorough and insightful analysis of law making within the United Kingdom in the period 1999-2004.

Throughout this book, it is possible to identify key themes such as the extent to which devolution has resulted in change through the transfer of powers to each devolved region, and extent that continuity exists within the law-making processes in the central UK institutions. The authors also consider whether it is possible to identify the establishment or development of distinct legal cultures within each of the devolved regions. Furthermore, the authors consider whether the emphasis placed by the new institutions on concepts such as transparency and accountability can be seen as a form of “new politics” when contrasted with the Westminster style of government.

When considering whether devolution has resulted in continuity or change, the authors have considered *inter alia* the approaches to law making within the devolved regions and the central UK government. For example, in his analysis of the role of Whitehall in law making since devolution, Trench argues that the changes to the machinery of government and the law making process have been “limited, minimal, ad hoc and incremental”. He asserts that there has been no major rethink of the British constitutional arrangements to accommodate devolution, and that consequently, devolution has meant only “limited constitutional change”. He continues that the failure to assign an individual or team to provide Whitehall departments with authoritative guidance on how “legislation should treat a devolved territory” and “to monitor the working of devolution, and in particular the way government departments seek to legislate” contributes to inefficiency. In their analysis of devolution in Northern Ireland, Morison and Anthony have also highlighted instances of continuity, particularly regarding the role of the Northern Ireland Civil Service in policy formation both under direct rule and during devolution.

Despite this, the authors did find it possible to highlight some changes that have been introduced, such as the innovations in legislative structures and processes that were adopted by the Scottish Parliament. In his discussion of the Scottish institutional arrangements, Page argues that in contrast to Westminster, the Scottish Parliament is not simply a mechanism for enacting the executive’s wishes into law, despite the dominance of the executive. When considering the Welsh Assembly, Rawlings opined that although the institution’s powers are limited to secondary legislation, its establishment

nonetheless “represents a quantum leap” in comparison to pre-existing arrangements.

Perhaps the area where there is the greatest potential for change under devolution is in the development of distinct legal cultures within the devolved regions. This is a complicated issue as there are “countervailing pressures for convergence and divergence”, with the former pushing for minimum standards across the UK, and the latter advocating unique policy initiatives to address the concerns of the populations of the devolved regions. In his discussion of the work of the Scottish Parliament, Page explains that some measures are introduced to replicate laws in England, and Sewel motions, which permit legislation from Westminster on devolved issues to apply to the relevant regions with the consent of the devolved institutions, are frequently employed. Subsequently, Rawlings asserts that much devolved Welsh legislation mirrors English laws. Nonetheless, both commentators have found evidence that distinct legal cultures are developing, as Page highlights the increased volume of legislation relating uniquely to Scotland that has been introduced since devolution and Rawlings asserts that it is possible to identify the growth of “Welsh legislative difference”. Similarly, Patchett asserts that there have been instances where the Welsh Assembly has declined to bring into force legislative provisions that were assigned to it. In relation to Northern Ireland, Morison and Anthony assert that, whilst in operation, the Northern Irish Assembly was very active and had “the potential for effective and vibrant law making”. The impact of Britain’s different legal traditions and the development of different legislative cultures in the resolution of devolution disputes have not yet been fully considered by either the executive or the courts according to Gee. He argues that these issues need to be integrated into the proposals for the establishment of a British Supreme Court, by, for example, ensuring the judges from each of the different legal traditions are represented on the bench.

It is also possible that devolution will be a catalyst for change within the central British constitutional arrangements, due to the emphasis within the devolved institutions on values such as transparency, participation and accountability. The impetus for the devolution project came *inter alia* from a sense that English concerns dominated law making within the United Kingdom, resulting in perceived disadvantages particularly concerning economic problems. It has been hoped that devolution will address these concerns by allowing decisions to be made “closer to the people”. The authors in this book have considered this issue by analysing the law making mechanisms established by devolved institutions, and the scrutiny at Westminster over legislation which affects the devolved regions. In his chapter on the Scottish Parliament, Page asserts the institution was specifically designed to be different to Westminster with more consensus, not confrontation between the political parties. It was hoped this would reduce majority rule and produce legislation that reflected the views of a larger proportion of the population. Subsequently, Rawlings asserts that the working practices of the Welsh Assembly represent “new politics” with greater provision being made for scrutiny of legislation, in addition to arrangements for equality of opportunity and bilingualism in the law-making

process. Subsequently, in their discussion of the Northern Irish devolved institutions, Morison and Anthony highlight the importance of the concepts of “new politics” in context of the conflict in Northern Ireland and argue that although these issues have not taken hold sufficiently to contribute to “truly wide-ranging and sustained local legislative preferences”, they are nonetheless reflected in the constitutional architecture and the tendency to conduct widespread consultation on an array of issues.

The level of scrutiny at Westminster over legislation affecting the devolved regions was considered by Winetrobe in his chapter on Sewel motions and law making for Scotland that is conducted in Westminster. He asserts that the employment of the Sewel Convention is problematic as “the parliamentary processes by which it operates are unsatisfactory, inappropriate, and inefficient, and could eventually undermine the legitimacy of the Convention itself, and even of the legislation produced through it”. He highlights that the processes used to scrutinise Sewel motions at Westminster are the same processes that led to devolution being deemed necessary, and that the process by which Holyrood approves a Sewel motion is not as rigorous as its scrutiny of executive policy proposals. In addition, he shows that those Scottish civil servants or ministers who will be responsible for the legislation’s implementation do not appear before Westminster’s Committees, and are therefore not held strictly to account. Similarly, Patchett in his consideration of how Westminster and Whitehall legislate for Wales, asserts that formally the Welsh Assembly is excluded from the process of making primary legislation, and therefore, Wales has less influence over laws that affect it than it did before devolution, although there have been some limited moves to allow the Assembly a voice during the deliberative stages. In addition, Hazell asserts that Westminster is able to “provide specialist territorial scrutiny of bills which apply solely to Scotland and Wales; less so for bills applying solely to Northern Ireland”, but that in “multi-territory” bills, English concerns tend to dominate.

This book provides thorough and up-to-date assessments by respected academics of the key issues relating to law making under devolution. Its target audience includes academics and students in law and politics, and government practitioners and policy makers. The analyses are generally clearly written, although some sections may be too complex for non-specialists. The structure of the book was clearly organised and the key issues flowed consistently through each of the chapters facilitating comparisons between the devolved institutions and central government. The authors also raise many interesting ideas for future consideration as the practice and impact of devolution become more entrenched.

Louise Mallinder
Queen’s University Belfast