# TABLE OF CONTENTS

**Editorial Note**  
RICHARD HANSTOCK  
4

**President’s Foreword**  
SHABANA SALEEM  
6

**Secret Trials and Special Advocates: The dangers of moving from the last resort**  
PAUL LIVINGSTON  
7

**Challenging Jurors**  
SARAH GOOD-ALLEN  
17

**Loss of Control: An Answer to Homosexual Discrimination?**  
LAUREN CROOK  
34

‘Live and Let Live’: The responsibility of women in the maternal-foetal conflict  
ELIZABETH REDINGTON  
65

**Mental illness and diversion from the criminal justice system**  
CELYN COOPER  
86

**Policing the Internet: The ‘Safer e-Communities’ Model**  
RICHARD HANSTOCK  
104

**Command Responsibility and Successor Liability**  
SARA FANTONI  
168

**To what extent are Economic, Social and Cultural Rights applicable in times of Armed Conflict?**  
MICHEÁL T MURPHY  
173

**Defending a General charged with war crimes: a summary of the case of Vinko Pandurević at the ICTY**  
HELENA TOŠIĆ  
190

**Indus Water Treaty in the ICJ: An Alternative Approach to Resolve Indo-Pak Water Disputes**  
AQEEL MALIK  
198

**Institutional Reform of Competition Law Authorities: Accounting for Psychological and Economic Factors that Influence Behaviour**  
SAM PRICE  
226

**To what extent have the cases of Viking, Laval and Demir led to change in the ability of trades union to take effective strike action in Europe?**  
ALEX JUST  
238

**Intellectual Property Law & the needs of consumers, creators / innovators and the public interest: An appropriate balance?**  
ANDREW PELLINGTON  
260
<table>
<thead>
<tr>
<th>Title</th>
<th>Author</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overriding Interests under the Land Registration Act 2002</td>
<td>MATTHEW ASTLEY</td>
<td>285</td>
</tr>
<tr>
<td>Directors, Codes of Practice, and the Companies Act 2006: Accountability Guaranteed?</td>
<td>ANDREW PELLINGTON</td>
<td>298</td>
</tr>
<tr>
<td>Risky Business: Challenges and Changes in the Assessment of the Standard of Care Owed to Participants in Supervised 'Extreme' Sports</td>
<td>FREDERICK LYON</td>
<td>316</td>
</tr>
<tr>
<td>Has Bolitho redressed the balance between defendants and claimants in cases of clinical negligence?</td>
<td>CELYN COOPER</td>
<td>332</td>
</tr>
<tr>
<td>The Human Tissue Act 2004</td>
<td>FIONA SHORE</td>
<td>351</td>
</tr>
<tr>
<td>ADR, UK Sport and the structure of Tribunals</td>
<td>DHARMINDER RUPRA</td>
<td>367</td>
</tr>
<tr>
<td>The Accentuated Advocate</td>
<td>CHARLOTTE DONALD</td>
<td>375</td>
</tr>
</tbody>
</table>
Editorial Note

RICHARD HANSTOCK

It is a privilege to have been entrusted with the compilation of this volume of the Gray's Inn Student Law Journal. Even though this Journal is still in its infancy, the response to my Call for Papers this year has been immense, and I have endeavoured to accommodate as many submissions as possible. The result is a diverse compilation of essays and articles, offering a cross-section of the interests and experience within the Gray's Inn student community.

This volume begins with a topical piece on closed material proceedings, moving into criminal procedure with an essay on the right to challenge. Articles on crime and criminology follow, including my own piece on the future of Internet policing. An international outlook continues with essays on successor liability, rights in times of armed conflict, and a summary of the case against General Pandurević from a member of his defence team. A new perspective on competition law is encouraged in a later article, followed by assessments of the right to strike, the state of intellectual property law and the accountability of company directors. Questions of medical ethics are considered, and gripping mountaineering stories punctuate an essay on negligence in extreme sports. This volume concludes with a special journalistic piece that is sure to turn a smile, on regional accents at the Bar.

As one might expect from aspiring barristers, each of these essays and articles are united by a concern for fairness and individual rights. This Journal reflects the interests of a lively community of current and recent students, many of whom will continue to play a part in the fabric of Gray's Inn for years to come.

Owing to the diversity of these publications, my editorial input into these essays has been limited. Full credit is due to the respective authors, whilst any errors in formatting are mine alone.
As well as being available electronically, this is the first volume of the Gray's Inn Student Law Journal to be committed to print. Copies will be placed in the Gray's Inn Library, and the electronic version will be available worldwide. It is my hope that as well as providing interesting and useful reading, sparking dialogue between students in a range of disciplines, this Journal might serve to attract future students to join the community that underlies this publication, and to get involved in the Committee responsible for arranging it.

None of this would have been possible without the support of the Education Department, and that of my peers on the Committee of the Association of Gray's Inn Students. I am particularly grateful to the volunteers who helped me to review submissions to this Journal, especially Matthew Astley and Justin Tadros. Though of course, the real stars of this publication are the dozens of students who took the time to submit their work for consideration. Thank you all.

RICHARD HANSTOCK

Suggested citation:
President’s Foreword

SHABANA SALEEM

It is with immeasurable pride that I write this foreword. It took a great deal of commitment and hard work to launch the Gray’s Inn Journal. In the words of Ben M’Hidi, “it is hard to start a revolution, but even harder to continue it”.

While the Gray’s Inn Journal may not have been ‘revolutionary’, it is without any doubt that I can say it has taken even harder work and commitment from the numerous contributors and the Journal Editor, Richard Hanstock, to revive this pioneering Journal.

Critics have often pointed to the difficulties of students producing and editing scholarly work. What skills can students offer to legal scholarship? How can students, so little experienced, contribute to the development of legal studies? The Journal contributors have proven that our students have an important role to play in legal scholarship and in informing legal practice.

I enjoyed reading every article and learnt a vast amount from overriding proprietary interests to the diversity of accents in advocates. I would like to thank each contributor, on behalf of AGIS, for not only taking the time to write but on a more personal note for educating me on such a variety of legal topics.

I would like to ask you to support the contributors by promoting the Journal in forwarding it to students, practitioners and academics. I hope you will find the articles as interesting as I have.

SHABANA SALEEM

---

1 Shabana Saleem, President of the Association of Gray’s Inn Students 2011–2012.
Secret Trials and Special Advocates: The dangers of moving from the last resort

PAUL LIVINGSTON

Introduction

Sent out for consultation to nestle quietly amongst the star-studded Leveson Inquiry and the widely-reported Legal Aid Bill, the ‘Justice and Security’ Green Paper¹ should in fact have been accompanied by red flashing lights. With an Opposition seemingly unwilling to make anything other than the most populist stands in the justice sphere, the right to a fair trial is likely to be dangerously undermined without so much as a whimper from our elected politicians. Although there have been excellent responses to this consultation from organisations, groups and individuals, the fact that these only number around 90 demonstrates a lack of engagement in an attempt to impose secret trials on an unsuspecting public.

The Government proposes to introduce legislation that will make ‘closed material procedures’ (CMPs) more broadly available in civil proceedings where sensitive material is relied upon. Material will be sensitive where, for example, disclosure of it may put an informant in danger or hinder the effective operation of the security services. CMPs allow the Government to avoid disclosing the ‘closed material’ to individuals or their legal representatives and to instead appoint ‘Special Advocates’. They are independent, security-cleared lawyers, who will have access to the closed material and have the responsibility of fighting for the maximum amount of information to be disclosed. Once the closed material has been served on the Special Advocate, they must then represent the individual to the best of their ability.

Lord Denning once espoused that “If the right to be heard is to be a real right which is worth anything, it must carry with it a right in the accused man to know the case which is made against him”² and CMPs are therefore a departure from this fundamental right. The most frequent use of CMPs has been in control order

---

¹ Justice and Security Green Paper, Cm 8194, October 2011 (‘Green Paper’).
or immigration appeal cases; cases that remain in the civil law sphere but which have an enormous effect on the lives and liberty of those concerned. Although the Supreme Court confirmed in *Al Rawi*\(^3\) that CMPs cannot be used in the absence of statutory authority, this Green Paper appears to be an attempt to provide sufficient statutory authority that CMPs will be possible in any civil case where the Government feels there is a public interest in non-disclosure.

This article will not reiterate the persuasive objections to these proposals that have already been made on the grounds that CMPs are contrary to Article 6 of the European Convention of Human Rights (ECHR) and that their very existence, never mind extension, constitute an unacceptable affront to natural and open justice. These arguments have been eloquently expressed elsewhere.\(^4\)

**The more evidence, the better?**

Firstly, I shall briefly examine one of the key arguments for these proposals; that it allows the court to consider all material rather than excluding some because it is sensitive and cannot be disclosed and that: “A judgment based on the full facts is more likely to secure justice than a judgment based only on a proportion of the relevant material”.\(^5\) Whilst this might sound superficially attractive on the basis that the more relevant evidence a judge has, the more informed a decision he can come to; on closer analysis it is both simplistic and deceptive. The truth of the matter is that evidence is only valuable when it can be properly challenged, hence the rules on hearsay evidence and the fundamental rights of a person to hear the case against them. Not only is such evidence sometimes worthless; it may also distort the facts in such a way that even an Independent judge will be unable to guarantee the fairness of proceedings. When a similar argument was previously made by the Government in *Al Rawi*, it was rejected by Lord Kerr.\(^6\)

---

\(^3\) [2011] UKSC 34.
\(^5\) Green Paper, paragraph 2.3.
\(^6\) *Al Rawi*, paragraph 93.
The effectiveness of the Special Advocate system

The main thrust of this article is to examine the very premise of the Green Paper; that: “The contexts in which CMPs are already used have proved that they are capable of delivering procedural fairness.” It will be submitted that this is a policy predicated on a fallacy, flying in the face of consistent criticism about the serious practical defects in the system from Special Advocates, the courts and academics. A typically punchy example is Lord Bingham’s comment in *Roberts v Parole Board*⁸ that a Special Advocate’s role would inevitably be “taking blind shots at a hidden target”.

Based on articles and evidence provided by Special Advocates and reports of the Joint Committee on Human Rights, there are at least four important hindrances on the ability of Special Advocates to effectively mitigate the unfairness suffered by individuals who cannot hear the case against them:

**(1) The prohibition on direct communication**

This represents the most significant restriction on the ability of Special Advocates to operate effectively. Although Special Advocates are, in severely restricted circumstances, allowed to communicate with the individual through the court after they’ve seen the closed material, this is very rare. Partly, this is because it is not considered tactically desirable for them to do so because they are required to notify the other side, leading to the risk that such an application might give away the parts of the closed material for which they have no explanation. They are also unlikely to be permitted to do so if the communication is anything to do with the closed material, making this practice almost entirely redundant.

---

⁷ Green Paper, paragraph 2.3.
⁸ [2005] UKHL 45.
This causes profound concern among Special Advocates as it prevents an individual from providing an exculpatory explanation for *prima facie* incriminating information. For example, evidence provided by an agent states that Mr X was at a meeting at a restaurant in Kabul at a specific time, but this allegation is not revealed to Mr X in the interests of national security. The Special Advocate has no way of putting this allegation to him, and even if Mr X would have been able to provide cast iron evidence that he was elsewhere, or that the person providing the secret evidence was prejudiced against him, he will be unable to do so.

Although it might be thought that the skills obtained in years of work as a barrister would adequately equip Special Advocate to avoid compromising secret information, the Government contends that without detailed knowledge of the investigation or other linked investigations, allowing communication could lead to inadvertent disclosure of sensitive information. They therefore assert that any proposed communication must always be reviewable by those familiar with the investigation and be cleared on the advice of relevant experts.

The current CMP regime that applies to Employment Tribunals is slightly more flexible, appearing to permit communication between Special Advocates and individuals that does not concern the substance of the closed material. However, the Green Paper seeks to abolish this flexibility in the name of harmonisation, on the basis that it will “bring a greater degree of consistency to proceedings in which Special Advocates are appointed.” Whilst this may indeed be correct, the argument for harmonising downwards rather than upwards is wholly unpersuasive.

A possible solution to this that has been raised by the Government is the use of a properly functioning ‘Chinese wall’ between Government counsel and those who

---

9 Suggested in *Forcese, C & Waldman. L*, ‘Seeking Justice in an Unfair Process, Lessons from Canada, the United Kingdom and New Zealand on the Use of Special Advocates in National Security Proceedings’ (Canadian Centre for Intelligence and Security Studies, August 2007).
10 Green Paper, paragraph 2.30.
11 See *Rahman v Metropolitan Police* ET Case Nos 220288/07 and 3200438/07.
12 Green Paper, paragraph 2.38.
would be reviewing the communications request.\(^{13}\) The idea behind this is to reduce the tactical disadvantage of requesting communication with the individual and therefore encourage the use of the existing procedures for communication. Whilst it is certainly welcome that the Government recognises the problem and is attempting to do something about it; it is far from a panacea. The Government admits there is a prevailing difficulty in regularly sourcing an independent team of officials to review the request and is also considering the “resource implications” of such a move, which are likely to be unpalatable in the midst of slashing justice budgets. These practical difficulties may well make this proposal unworkable.

Special Advocates have provided two proposals with regard to the bar on direct communication, contending that even these are “relatively modest by comparison with the more relaxed treatment of closed evidence in procedures in other jurisdictions”:\(^{14}\)

(i) At the very least, communication on matters not related to the substance of the closed procedure should be permitted in all CMPs. This would represent an upwards harmonisation to the current position in Employment Tribunals.

(ii) Special Advocates should be allowed to apply \textit{ex parte} to the Court for permission to communicate about the substance of the closed material. Where there is no question of sensitive disclosure, the Court should be able to authorise communication without reference to the Secretary of State. Where there is doubt, the Special Advocate should be able to pursue the application whilst giving notice to the Secretary of State, or abandon it.

Furthermore, the model used in the United States, where directly instructed representatives (rather than state appointed Special Advocates) are provided with closed material subject to a ‘Protective Order’ is raised and Special Advocates have asked for a reasoned justification as to why this could not be

\[^{13}\text{Green Paper, paragraph 2.33.}\]
\[^{14}\text{Justice and Security Green Paper, Response to Consultation from Special Advocates (16 December 2011), paragraph 31.}\]
adopted in the UK. Such proposals represent an attempt to properly balance the right to a fair trial and the need for national security and should be seriously considered as they come from those at the very heart of the process.

(2) The lack of access to independent expertise and evidence and the consequent limitations on arguing for disclosure

Although Special Advocates are technically allowed to adduce expert evidence, their practical inability to do so can be regarded as one of the principal imbalances in practice when attempting to challenge the Government’s case or request greater disclosure.

The difficulty is that there is nobody that Special Advocates can call to provide independent expertise and evidence. Potential persons would have to have the requisite expertise in the area (e.g. foreign terrorism, national security or domestic intelligence), recent inside knowledge and some element of independence. Although an obvious category of potential experts would be recently retired security or diplomatic personnel, they are either unwilling or unable to give evidence against their former employer; meaning that no Special Advocate has ever been in a position to adduce expert evidence.

Where an allegation is made against an individual that is largely based on closed material, this is where Special Advocates should be at their most strident in arguing for greater disclosure, yet in practice is where they are most impotent. This is particularly worrying with the proposed expansion in potential use of CMPs – it means that in any civil proceedings in which the Government is concerned about disclosing information, it can assert that secrecy is necessary in the public interest and this will be effectively unchallengeable.

A case that would be handled differently under the new proposals is the recent litigation involving Binyam Mohammed, where the Government chose to settle the case rather than disclose seven subparagraphs of evidence, as it was claimed

\[15\] Ibid, paragraph 32.
that disclosure it would lead to a real risk of serious harm to the national security of the UK. Under the Green Paper proposals, the Secretary of State would apply for this to be heard under a CMP, with a Special Advocate attempting to argue for disclosure without being able to adduce any evidence whatsoever that could dispute the Government’s evidence for non-disclosure. They will be trying to shoot a gun without any bullets; and as Angus McCullough (a current Special Advocate) points out: “Courts inevitably accord great weight to views on matters of national security expressed by the agencies who are particularly charged with protecting national security”.16

This inequality of arms places a serious limitation on the Special Advocates’ ability to fulfil their functions effectively, and I would suggest that this should not be compounded by extending the use of Special Advocates beyond their current spheres. Although the Government proposes to ameliorate this by greater training of Special Advocates, this is unlikely to seriously improve their ability to gainsay the evidence of the Security Services.

(3) Tainting

The problem of ‘tainting’ is similar to difficulties of No 1 above, but restricts communication before the closed material has been served. The concept is that once a Special Advocate has seen closed material in one case, he is potentially ‘tainted’ because of the knowledge that he has gained and is therefore unable to communicate with individuals in certain subsequent cases. In practice, this occurs if the Special Advocate has previously acted in any case involving secret evidence from the same country or one with any factual connections; therefore the more experienced a Special Advocate is, the worse position in which he will find himself.

The complete lack of communication is an obvious derogation from the normal lawyer-client relationship and makes establishing any relationship with the

individual almost impossible, as even basic questions cannot be asked. It could be eased if the Government were more prepared to trust the judgment and discretion of the Special Advocates, as well as taking a more realistic view of their likely inability to recall details of previous cases, however this seems unlikely and is not mentioned in the Green Paper.

(4) Late disclosure of evidence

Several Special Advocates have also pointed out the ‘endemic’ problem of very late disclosure of documents by the Government and the serious concern that this causes. Although in normal cases the court can simply refuse to admit late evidence or allows an adjournment, this is understandably unlikely in national security cases which have to proceed on full facts and as quickly as possible. Furthermore, cost sanctions are inappropriate as all parties are generally publicly funded.

A lack of proper sanction against this creates a serious disjuncture between the supposed role of a Special Advocate and that which they are able to perform in practice. At present, very late disclosure of evidence means that the Special Advocate’s role of arguing for the disclosure of the closed material effectively “goes out the window”.17 The Government’s rejection of this as a problem and statement that the judge can adjourn proceedings if they feel that prejudice has been caused18 smacks of them putting their head in the sand. One possible solution raised by Special Advocates has been the imposition of more active case management powers, insofar as they involve some enforceable and enforced requirement on the Government to comply with court imposed deadlines.19 Any such powers would be welcome as at least an attempt to stem this systematic unfairness.

PAUL LIVINGSTON

17 JCHR 2010, Evidence Question 99.
18 Green Paper, Appendix F.
19 Special Advocates Response, paragraph 35.
Conclusion

The proposed extension of the use of CMPs has been barely reported in the press and appears to have little opposition inside Parliament. This is a huge mistake, as Lord Kerr stated in *Al Rawi* that CMPs should: “Always be a measure of last resort; one to which recourse is had only when no possible alternative is available. It should never be regarded as an acceptable substitute for the compromise of a fundamental right”.\(^{20}\) These proposals constitute a transformation of a system used as a last resort by the Courts in exceptional circumstances, to one where the Secretary of State can enshrine a case in secrecy as soon as he thinks disclosure of sensitive material would cause “damage to the public interest”.\(^{21}\)

The Government clearly admits in the Green Paper that: “How well the Special Advocate system works will be a critical factor in the success of the proposed expansion of CMPs into new contexts.”\(^{22}\) They are absolutely correct in this regard, yet as has been shown above, a system that is riddled with practical defects in the ability of Special Advocates to properly represent individuals who are excluded from the proceedings against them is not one that is fit for purpose.

The Green Paper is based on the faulty foundations of a Special Advocate system that has cracks all over it. The practical bars on direct communication makes the task of defending an individual against incriminating information extremely difficult, whilst the lack of access to independent expertise ensures that Government assertions about the necessity of non-disclosure are effectively unchallengeable. Finally, very late disclosure of evidence and the Government’s unwillingness to recognise this as a problem can make the Special Advocate’s role in arguing for disclosure essentially non-existent. Expanding the use of CMPs into the wider sphere of civil law is an almost unprecedented attack on the fundamental principle of open justice and must be strongly opposed.

\(^{20}\) *Al Rawi*, paragraph 94.
\(^{21}\) Green Paper, paragraph 2.7.
\(^{22}\) Green Paper, paragraph 2.24.
The following are also useful sources


**Challenging Jurors**

SARAH GOOD-ALLEN

Impressed with the delivery of prompt justice following the riots of summer 2011, government ministers are now considering eliminating the right to trial by jury for less serious offences; a proposal that, if implemented, would be considered by many as a removal of one of the fundamental rights to a fair trial. However it is curious to ponder over why juries are inherently just. How does the constitution of, and the challenges to, juries make them impartial? Are the methods of challenge to jurors that are open to prosecution and defence counsel fair and do they lead to juries who are able to try cases justly? This paper aims to shed some light on this conundrum.

Juries have been used, albeit for different functions from the present day equivalent, since the Norman Conquest\(^1\). Over time the jury has taken on a more judicial role; this has been particularly prevalent since the Magna Carta in 1215, recognising a person’s right to be trialled by a jury, and notably since *Bushell’s case*\(^2\), which provided one of the most important vehicles for the development of an independent jury\(^3\). In the present day, juries are most commonly used in criminal cases; according to Glendon et al\(^4\) less than 1% of civil trials involve juries, and of criminal cases, only a small (1%), yet by no means trivial\(^5\) quantity remain to be decided by a jury.\(^6\)

Currently, the law takes steps to minimise the risk of prejudice or bias of a jury. Both Lord Hope of Craighead in *Porter v Magill*\(^7\) and Lord Rodger of Earlsferry, dissenting, in *R v Abdrolkov*\(^8\) demonstrated that there are safeguards in place to assist in reaching an impartial and just verdict. The European Court has also

---

2. (1670) 124 ER 1006.
ruled that sufficient guarantees must exist to exclude any objectively justified or legitimate doubts as to the impartiality of the jury, and stressed that a jury must be impartial from a subjective as well as an objective point of view.9

Within the text, these ideas will be developed and incorporated with legislation and further case law in order to analyse how fair challenges to jurors are and if they lead to juries who are able to try cases justly.

Primarily, it is prudent to consider the constitution of juries and the legislation that governs their formation. The Juries Act 1974, which consolidates certain enactments relating to juries, jurors and jury service, has been amended by the Criminal Justice Act 2003, which altered, amongst other things, who is entitled to sit on a jury.

Section 1 of the Juries Act 1974 established the qualification for jury service. It specifies that jurors are drawn from the electoral register and, of those people, anyone aged between 18 and 70 can be summoned for jury service, so long as they do not suffer from a mental disorder, nor are they disqualified for whatever reason. Following amendments made by the Criminal Justice Act 2003,10 “members of the judiciary, lawyers, those concerned with the administration of justice, the clergy...those aged 65 or over, those who had served on a jury full time serving members of the armed forces, and medical personnel”11 are now eligible to sit on a jury, whereas prior to 2004, when the Criminal Justice Act 2003 came into force, they were either ineligible or they possessed an excusable right from service. It is now only those people with mental disorders12 13 or with criminal convictions14 who remain ineligible for jury service. These changes, proposed by the Auld Review,15 were designed to yield “fairer and more effective

---

9 Sander v United Kingdom (2001) 31 EHRR 44.
10 Criminal Justice Act 2003 Sch 33.
12 Juries Act 1974 Sch 1 Pt 1.
14 Juries Act 1974 Sch 1 Pt 2.
trials”.[16] An example of this could be that, as members of the armed forces are now eligible for jury service, they may be able to provide greater understanding to the circumstances surrounding a case involving ex-military personnel. Whether this is a fair reflection and has indeed led to fairer, more effective trials will be seen later.

Despite the now greater pool from which jurors can be selected, there are still inadequacies in the selection process, which should be addressed when considering who can sit on a jury. Jurors are drawn from the electoral register[17] in order to represent a cross section of society drawn at random,[18] yet it is widely known that not everyone is included on the electoral register, particularly the young and ethnic minorities.[19] Consequently, a true reflection is not provided of those potentially eligible to serve and this decreases the likelihood of a representative sample of the population.[20] Moreover, research by Darbyshire[21] demonstrated that a high proportion successfully evade or avoid jury service.

Therefore, even though there are now more eligible people to serve, due to the reasons highlighted above this does not necessarily lead to a more representative and fairer jury panel.

Under the new eligibility rules, as set out above, police officers, lawyers and judges may serve as jurors. Appearance of these members as jurors on cases is tightly governed by guidelines issued by the Central Summoning Bureau,[22] to

ensure that they do not unduly influence their fellow jurors.23 Yet, despite these guidelines, cases such as R v Abdroikov24 and R v Khan25 are still appealed on the grounds that a member of the jury had an appearance of bias by reason of their occupation. These were both appeal cases that questioned whether, with regard to the fact that Parliament had declared that such persons were allowed to sit on juries, there was a real possibility that the jurors were biased.26 The House of Lords and Court of Appeal respectively, were both required to determine the potential influence of jury members on other jurors after the trials had been heard. However, there are methods used for challenging jurors prior to the trial being heard and these will be set out below.

Challenges to jurors – Challenging the array

The Juries Act 1974 confers to the prosecution and defence the right to challenge jurors. Challenges may be made against the jury as a whole, known as ‘challenging the array’ under s 5(2) of the Juries Act 1974. The defendant has the opportunity to challenge the way in which the jury panel was selected by being “entitled to reasonable facilities for inspecting the panel from which the jurors are or will be drawn”,27 on the basis that the selection process was unrepresentative or biased. In R v Danvers28 an application was made to challenge the array at the trial of a West-Indian defendant when all members of the jury panel were white. The application failed as it contained no allegation that the all-white jury had been empanelled as a result of bias. Currently, this type of challenge is extremely rare.

Indeed, challenges may be made to individual jurors by both the prosecution and defence by ‘challenging for cause’, as set out by s 12 of the Juries Act 1974, on the grounds that the would-be juror may be biased for or against the accused. Furthermore, the right to an additional challenge is held only by the prosecution

26 [2007] 1 WLR 2679 (HL) 2682.
27 Juries Act 1974 s 5(2).
and judge, whereby they can make a potential juror stand by and consequently prevent him from sitting on that jury. However, following the Attorney-General’s guidelines\(^29\) this is rarely used in practice.

Initially, challenging for cause and the merits offered to both prosecution and defence by this challenge will be considered, followed by the prosecution’s and judge’s right of stand by and whether this challenge is fair to both prosecution and defence.

**Challenging for cause**

Challenging for cause challenges the right of an individual juror to sit on the jury\(^30\) and, following a successful challenge, the potential juror cannot try the present case but may, so long as he is still regarded as a suitable juror, try another case.

The prosecution or defence may challenge for cause as many jurors as they deem necessary on the grounds that the juror is:

- Ineligible or disqualified; or
- Reasonably suspected of being biased.\(^31\)

When establishing this challenge it is the party who alleges the bias that bears the burden of proving that is so. This evidential burden is so high, as will be shown later, that it prevents both the prosecution and defence from using this challenge indiscriminately. In order to establish a challenge for cause, counsel says ‘challenge’ before the would-be juror takes his oath. The case of *R v Chandler*\(^32\) established that evidence substantiating the challenge must be provided before the juror may be asked any questions regarding the matter, placing a burden on the counsel preventing them from abusing this challenge. Furthermore, challenging for cause, and in particular questioning members of

---


\(^32\) [1964] 2 WLR 689 (CA) para 330.
the jury about potential bias, may only take place in exceptional circumstances, as illustrated in the cases of *R v Kray*[^33] and *R v Andrews*[^34] in which the Court of Appeal held that the questioning of jurors may be counterproductive and should therefore only be done in the most exceptional circumstances. It was during the trial of the Kray twins that the judge was prepared to exclude any members of the jury who had read some of the sensational news reporting.[^35] In *R v Andrews*,[^36] this point was reinforced with regard to the test for apparent bias in *R v Gough*,[^37] which set out that the court should ask itself,

> “whether, having regard to those circumstances, there was a real danger of bias on the part of the relevant member of the tribunal in question, in the sense that he might unfairly regard with favour, or disfavour, the case of a party to the issue under consideration by him.”

Therefore, if a member of the jury was so biased as to create 'a real danger' that he might unfairly regard the case which he was considering, then it would be right to challenge and even exclude this member.

This test was criticised in *Porter v Magill*[^38] for placing too much emphasis on the court's subjective view of the facts. In the Privy Council’s decision in *Montgomery v HM Advocate; Coulter v HM Advocate*,[^39] which was only persuasive rather than binding, it was held that the court was entitled to expect that “the jury will follow the instructions which they receive from the trial judge and that they will return a true verdict in accordance with the evidence”,[^40] thereby rendering the defendant's trial fair.

Through consideration of these cases and that of *Re Medicaments and Related Classes of Goods (No 2)*[^41] it was held in *Porter v Magill*[^42] that a modest adjustment of the test was required, so that it no longer included the reference to 'a real

[^34]: [1999] Crim LR 156.
[^40]: [2003] 1 AC 641 (PC) 674.
[^41]: [2001] 1 WLR 700 (CA) 711.
[^42]: [2001] UKHL 67 (HL) para 103.
danger” but instead, ‘a real possibility’. As Lord Hope noted,\(^43\) this produces an objective test of the reasonable man, which is in line with the jurisprudence of the Strasbourg court and therefore is fairer than the subjective test outlined in *R v Andrews*.\(^44\)

It appears that subsumed within this test is the ubiquitous notion that justice must not only be done, but must be seen to be done; a defendant must receive a fair trial and one way that this can be ensured is through challenging jurors for cause to eliminate potential bias.\(^45\)

In slight contrast to this, the Court of Appeal acknowledged an important point in *R v Abu Hamza*\(^46\) - that the risk that a jury may be affected by prejudice is not one that can be wholly eliminated. Subsequently, due to human nature there may be a natural element of bias that cannot be wholly removed despite the use of challenges.

This challenge may be regarded as fair to both prosecution and defence as it is available to both parties and requires them to fulfil the same criteria. However not all challenges to juries may be considered as fair as this.

**The prosecution’s right of stand-by**

The defence, prior to 1988, possessed a right to challenge potential jurors ‘peremptorily’ without providing a reason, but this has since been abolished owing to the fact that is has been viewed as a waste of time and a corruption of the judicial process. Interestingly though, the prosecution have retained this right, known as ‘the prosecution stand-by’.\(^47\)\(^48\) When exercising this right the prosecution must say ‘stand by’ just before the juror takes his oath. In this

\(^{43}\)Ibid.

\(^{44}\)[1999] Crim LR 156.


\(^{46}\)[2006] EWCA Crim 2918.


\(^{48}\)Criminal Justice Act 1988 s 118(1).
instance a juror may be allowed to sit on another case, but will not be able to sit on that jury.

The right to stand by is rarely used following the guidelines set out by the Attorney-General in 'The Exercise by the Crown of its Right to Stand By'\(^{49}\) and moreover, as it may be seen as unfair to extend this right to the prosecution and not the defence,\(^{50}\) further guidance is provided to Crown Prosecutors.\(^{51}\) This challenge may be regarded as giving the prosecution excessive amounts of power, over and above the challenge available to the defence, by providing them with the advantage of being able to effectively ‘rig’ the jury. Hence, the Crown Prosecution Service have outlined the circumstances in which it would be proper for the Crown to exercise its right to stand by a member of the jury panel; they are as follows:

- “to remove a manifestly unsuitable juror, but only if the defence agree; and
- to remove a juror in a terrorist or security case in which the Attorney General has authorised a check of the jury list, but only on the authority of the Attorney General.”\(^{52}\)

Clearly the prosecution’s right of challenge is greatly restricted as it would only be appropriate to remove a ‘manifestly unsuitable juror’, not merely \textit{any} juror, and it is coupled with the further element of agreement of the defence. This in some way negates the fact that the defence do not have this type of challenge as their agreement in removing a juror is required. It is this author’s view that despite the right of stand-by being seen as manifestly prejudicial, one-sided as it is only available to the prosecution and a possible corruption of the judicial process, it is necessary and fair. This is due to the fact that it eliminates any potential bias of those challenged and its use is closely regulated by two sets of


\(^{50}\) P Hungerford-Welch \textit{Criminal Procedure and Sentencing} (7th edn Routledge-Cavendish, Oxon 2009) 519.


guidelines, both of which have high standards to satisfy. Furthermore, the defence must be party to a decision by the prosecution to remove a juror, demonstrating that this challenge is not entirely one-sided in favour of the prosecution.

The judge’s right of stand-by

The judge also has an inherent right to stand a juror by, but this, as with most of the other challenges to jurors, is used sparingly.53 Lord Lane outlined in the case of *R v Ford*,54 that a judge’s right of stand by “is to be exercised to prevent individual jurors who are not competent from serving. It has never been held...to influence the overall composition of the jury”55 thereby ensuring that fairness is achieved by the principle of random selection. The essence of this decision was reiterated in the case of *R v Tarrant*,56 where the Court of Appeal held that a judge should not use his discretion to discharge individual jurors in order to alter the composition of the jury panel.

Interestingly, the case of *R v Smith*57 indicates a possible scenario where it may be possible for a judge to take steps to interfere with the jury’s composition. In that case the judge did not accept that the "trial would only be fair if members of the defendant’s race were present on the jury".58 However this observation suggests the possibility of a judge including ethnic minority representation on the jury if the case were one that required knowledge specific to that ethnicity. Davies and Edward’s59 research supports the idea that this possibility would produce a fairer result. They extend their argument further by theorising that if a jury comprising some members of the defendant’s race is likely to bring about understanding of the defendant’s background, thereby producing a fairer result, perhaps this is the best method of ensuring an impartial jury, and not random

57 [2003] EWCA Crim 283.
58 [2003] EWCA Crim 283 (CA) para 40.
selection. This raises concern surrounding the representativeness of juries and calls into question the issue of whether they are therefore fair within that regard. The Auld Review\(^6^0\) suggested that a provision be made that would account for such a scenario, increasing the fairness of any jury panel. When the Criminal Justice Act 2003 came into force however, there was no such provision; perhaps it is necessary for there to be one.

This point raises the question: do the challenges to juries enable them to try cases justly? In order to answer such a question a definition of ‘justly’ must be sought. Legal dictionaries\(^6^1\) suggest it means ‘impartial’ and ‘having a basis conforming to fact and reason’. Impartiality is regarded as subjective and, as Lord Hope in \textit{Porter v Magill}\(^6^2\) notes, is therefore very difficult to establish. If that is so, then there is an inherent difficulty in assessing whether challenges to juries lead to those who can try cases justly.

If however, these challenges are compared to those of other jurisdictions they may provide a more factual, objective basis upon which to analyse this question. In the United States multiple extensive checks are carried out on potential jurors in the hope of eliminating bias. Cook and Sullivan\(^6^3\) believed that the American citizen’s right to such extensive challenges would eliminate those jurors most biased, least responsible and least fair, and without these checks these members would be allowed on juries lessening the fairness of the defendant’s trial. Norton et al\(^6^4\) also noted that the use of peremptory challenge, despite allowing any criteria to be used in excluding jurors, could lead to a fairer trial for the defendant.

---

\(^6^1\) Legal-Dictionary <legal-dictionary.thefreedictionary.com> accessed on 15 April 2011.
\(^6^2\) [2001] UKHL 67 (HL) para 88.
In comparison, the right of jury challenge is very limited in England and Wales as illustrated in the Auld Review. The right of jury challenge is very limited in England and Wales as illustrated in the Auld Review. One reason for this came from the Roskill Committee’s Report which highlighted that extra challenges, in particular the peremptory challenge, eroded the principle of random selection. It is argued that random selection itself eliminates bias, as by having a jury panel made up of jurors with varied backgrounds, each juror’s bias cancels out the others. Furthermore, as previously mentioned, the elements required to, or the restrictions that must be adhered to, in order to secure a successful challenge are so high that the challenges themselves are rarely used. Auld also rightly points out that there is little support for altering the method of challenging jurors, suggesting faith of the Government, judges and the public in how juries currently try cases. Moreover, it is necessary to emphasise that the relative difficulty in challenging jurors should be viewed in light of England and Wales’ provision for majority verdicts, which in the US is unusual if not unheard of. This in itself “prevents individual jurors thwarting the otherwise general consensus of the jury” and consequently minimises any existing bias which the use of challenges, or lack thereof, may not have.

Jury research

When considering whether challenges to jurors lead to juries who can try cases justly, it is prudent to acknowledge the lack of research into juries in addition to the sacrosanct nature of the reasons for the jury’s decision. Thomas argues, in addition to Auld, that knowing of the juries reasoning may assist in assessing

---

68 ibid.
70 Contempt of Court Act 1981 s 8.
how useful a jury is in delivering a fair verdict. Padfield\(^73\) also suggests that not knowing of the reasoning has the potential to undermine the defendant’s right to a fair trial. This is demonstrated by the cases of \(R v \ Mirza,\)^74 in which a juror had suggested that there was an element of racial bias in the jury’s verdict, and \(R v \ Smith\)^75 where a juror told the judge of impropriety within the jury. Had these circumstances in both cases not been brought to light it would not have become clear that the jury’s reasoning was biased, and would not have emphasised how biased reasoning could undermine the defendant’s trial. Moreover, following the case of Attorney-General \(v\) Scotcher,\(^76\) the government decided that combating matters of impropriety should develop through the common law on a case-by-case basis. Padfield\(^77\) questions whether this is a satisfactory response, concluding that most research suggests no other system is better. This view is supported by the Auld Review\(^78\) in paragraph 74 where it is acknowledged that many England and Wales judges “would stand firm against any such attempts” at changing the way juries are challenged. Consequently, this current scenario may be regarded not as the ideal solution, but as the most appropriate. As alterations to the challenges of juries will not be supported, faith must be placed in the common law to develop ways of overcoming bias in juries. It is this author’s view that in order for case law to render an impartial jury, the government too must take steps to ensure that these developments are supported by legislation.

In order to determine whether cases have been tried justly because of the challenges of jurors, research needs to ascertain how jurors reached their decisions. Cases such as \(Mirza\)^79 and \(Smith\)^80 signify that unjust jurors do exist and are members of jury panels, demonstrating that challenges had not been made against those jurors. If challenges had been used, would they have


\(^{74}\) [2004] 1 AC 1118.

\(^{75}\) [2005] 1 WLR 704.

\(^{76}\) [2005] UKHL 36.


\(^{79}\) [2004] 1 AC 1118.

\(^{80}\) [2005] 1 WLR 704.
eliminated this risk of bias? It appears that this is difficult to tell through lack of research⁸¹ and due to the subjective nature of the answers that jurors give when asked of their opinions because, as Thomas⁸² notes, individuals may change their answers or opinions unintentionally, therefore giving no accurate representation of the individual’s conduct.

Conclusion
Within this paper four essential points have been discussed in relation to the main issues around the selection and potential influences on the performance of a jury within England and Wales. Since the changes brought in by the Criminal Justice Act 2003, a greater number of citizens are eligible to be called for jury service, thereby creating a more representative jury.

Jurors may be challenged by three prominent methods; challenging the array, challenging for cause and the prosecution’s and judge’s right of stand-by. Challenging for cause provides an extremely versatile and fair method of carrying out checks on jurors as, being open to both counsels, it provides a ‘level playing field’ upon which to challenge jurors. Furthermore, the evidential basis upon which a challenge can be brought is so high that it prevents the use of any unfettered discretion, meaning it remains a fair challenge.

The right of stand-by however, may not at first appear to be fair. In fact, it may seem to be indiscriminate, as the prosecution may challenge any member of the jury he deems necessary. Yet the restrictions imposed on the implementation of this challenge curb any potential bias resulting from its use. This is not only fairer on the defence counsel, who lacks this method of challenge, but it also avoids affecting the composition of the jury in a way that may result in it becoming less representative.

Due to the nature of these challenges it would seem as though they do indeed go some way to producing juries who can try cases more justly than those which are not challenged at all as, when successful, they eliminate jurors with the potential for bias. However, as demonstrated by *R v Chandler*\(^8^3\) and *R v Andrews*\(^8^4\) this rarely happens as the evidential burden is so great, ensuring that the principle of random selection remains intact.\(^8^5\) Indeed, the suggestion of adopting an approach such as that of the US would gain little support as, despite having a greater number of less restrictive challenges, it too would be regarded as undermining the desired representative outcome of the selection process of jurors. Perhaps then it is the method of selection and the need for greater information on jury panels that is crucial to yielding juries who can try cases more justly as Davies and Edwards\(^8^6\) suggest. By interfering as little as possible with the selection process, the outcome should be an impartial panel of jurors. As noted earlier, the increased number of people who are eligible for jury service has inevitably contributed to establishing more representative juries,\(^8^7\) as this is inherent in the nature of having a greater number of citizens to select from. However it should be reiterated that many individuals evade or avoid jury service and consequently, the representativeness of a panel may be diminished.\(^8^8\)

With regard to this paper, it can be seen that challenges to jurors and juries do lead to panels that can try cases more justly, but this is only to the extent that the principle of random selection is not eroded.\(^8^9\) *How* justly a jury can try a case will consequently be determined by the method of jury selection and greater research into how juries conduct themselves.

\(^8^3\) [1964] 2 WLR 689.  
\(^8^4\) [1999] Crim LR 156.  
\(^8^5\) *R v Tarrant* [1998] Crim LR 342.  
\(^8^7\) P Hungerford-Welch *Criminal Procedure and Sentencing* (7th edn Routledge-Cavendish, Oxon 2009) 505.  
\(^8^9\) *R v Ford* [1989] QB 868.
It must be accepted then, that due to the subjective nature of jurors and the lack of accurate empirical research on just trials, there will always be an element of bias that cannot be removed. Consequently, this paper suggests that challenges may lead to a more just panel, but not necessarily one that is wholly just.

SARAH GOOD-ALLEN

Reference List

Cases
Bushell’s case (1670) 124 ER 1006.
Montgomery v HM Advocate; Coulter v HM Advocate [2003] 1 AC 461.
R v Alan I [2007] EWCA Crim 2999.
R v Khan [2008] EWCA Crim 531.
R v Mirza [2004] 1 AC 1118.
Re Medicaments and Related Classes of Goods (No 2) [2001] 1 WLR 700.
Sander v United Kingdom (2001) 31 EHRR 44.

Legislation
Contempt of Court Act 1981

Criminal Justice Act 1988
Criminal Justice Act 2003
Juries Act 1974

Books

Journals

Web Resources

Bibliography
Loss of Control: An Answer to Homosexual Discrimination?

LAUREN CROOK

Chapter 1: An Introduction to the Issues

1.1. Domestic Violence in Homosexual Relationships

Unfortunately, for many, the home is no longer a safe haven and domestic violence has become a major social problem which raises a host of legal issues, especially surrounding the law of provocation. The range of abuse experienced through domestic violence is vast, and can include physical or psychological harm, rape or deprivation. Currently, the Government acknowledges domestic violence “regardless of gender or sexuality.”\(^1\) It is asserted that under a recurrence of these circumstances, both women and men irrespective of sexual orientation could be provoked into killing their abuser and thus, the law on homicide should protect all victims without favour to stereotypical heterosexual males.

More often than not, domestic violence is associated with the assertion of power by males over their female partners, completely disregarding homosexual relationships. Domestic violence should not, in the eyes of the law, define itself against what old society views as a ‘traditional relationship.’\(^2\) Schneider, a feminist, believes that domestic violence arises due to an abuse of power by a dominating male who believes that he has a sense of entitlement and sexual proprietariness.\(^3\) While it may traditionally correct to talk about ‘battered wives’, this is no longer the case.\(^4\) It is instead suggested that domestic violence arises out of dysfunctional families, where domestic violence is a mutual conduct between partners. Johnson labels this as “situational couple violence,” where violence is situationally provoked as emotions and tensions rise within a

---


This idea is supported by Home Office statistics that signify that 20% of men in England and Wales are also victims of domestic violence. This is due to the rise of acceptance of homosexual relationships, and the rise of domestic violence within these relationships.

Furthermore, studies have revealed that violence does occur in homosexual relationships. Henderson’s study of domestic violence among lesbians and gay men identified that 29% of the male sample questioned had suffered physical, sexual or mental abuse from a partner of the same sex. Similarly, 22% of the female sample suffered the same fate. Henderson went on to reveal that of the men who had experienced abuse at the hands of a regular male partner, 12% were suffering repeated abuse from their partner. The statistics also highlight that domestic violence is a prominent issue in gay relationships as it has been argued that it occurs at the same rate as violence in heterosexual relationships.

Positively, the increased awareness of the nature and severity of domestic violence in recent years has encouraged a change in attitude to view the domestic context of violence as an “aggravating rather than a mitigating factor.”

The most significant factor demonstrating seriousness of provocation is the objectively judged gravity of the provocation. Whether the cumulative effect of on-going abuse and multiple incidences of provocation should bear more or less weight is uncertain. However it is an important issue in the domestic context as the gravity of the last provoking event is often only understandable in the light of a background of incidents. While the courts have gradually begun to understand the full effect on systematic abuse, this has only been in a heterosexual context.

---

1.2. The Distinction Between Homosexuals and Heterosexuals in the Criminal Law

Legislation such as the Civil Partnership Act 2004 recognises that homosexuals should have equal rights.\(^{11}\) Similarly, the Domestic Violence, Crime and Victims Act 2004 recognises that same-sex couples can encounter domestic violence by virtue of cohabitation.\(^{12}\) The Act was introduced to ensure that domestic violence is treated with the same weight as all other crimes. Thus, the Act amends the Family Law Act 1996 to make a breach of non-molestation orders a criminal offence, and extends the provisions of the Family Law Act to same-sex couples.\(^{13}\) However unsatisfactorily, in general the courts are unwilling to view homosexuals on equal footing to heterosexuals. This was first demonstrated in the Report of the Committee on Homosexual Offences and Prostitution in 1957 and subsequent debate by Devlin and Hart, and secondly, in the distinction drawn between Brown and Wilson.

The Report of the Committee on Homosexual Offences and Prostitution in 1957, otherwise known as the Wolfenden Report, sought to decriminalise homosexual practice between consenting adults in private. The Committee argued that no act of immorality should be made a criminal offence.\(^{14}\) Whilst the Committee supported the notion that morals should not be enforced through the law, it is clear from the wording of the report that homosexuality was still viewed as immoral. Hart supported the report, stating that the law has no business to control and punish activities enjoyed in the private realm.\(^{15}\) On the other hand Devlin contended that certain practices are morally wrong, and it is the duty of the law to prohibit such actions, regardless of that fact that it lies within the private realm.\(^{16}\) The recommendations led to the passage of the Sexual Offences  

\(^{11}\) Civil Partnership Act 2004.  
\(^{12}\) Domestic Violence, Crime and Victims Act 2004 s 3.  
Act 1967, which legalised homosexuality.\textsuperscript{17} What is unfortunate about this debate is that whilst Hart expressed the more liberal view that the law should not constrain homosexual practice, at no point did he dispute that homosexuality is immoral, he merely expressed that to enforce morality through the law is unjust.\textsuperscript{18} It can be seen from this historical background to the law that the law has been built by judges who have a low or no tolerance for homosexuality. This of course has a great impact on the application of the law, which is demonstrated in the following cases.

In \textit{Brown} the House of Lords (HL) were obligated to determine the legality of sado-masochistic activity between consenting homosexuals.\textsuperscript{19} It was held that such activities were unlawful. Rejecting arguments that a free society must tolerate these activities, the majority concluded that freedom of expression is not enough to justify the infliction of intentional injury. Lord Jauncey of Tullichettle expressed that “it would not be in the public interest that deliberate infliction of actual bodily harm during the course of homosexual sado-masochistic activities should be held to be lawful.”\textsuperscript{20} The latter case of \textit{Wilson} involved a married man who branded his wife on the buttocks, at her suggestion, using a hot knife. The Court of Appeal (CA) upheld the appeal, declaring that it is contrary to public interest for the state to interfere in the private activities of husbands and wives.\textsuperscript{21}

Whilst \textit{Wilson} upholds that sexual freedom involving consensual violence is a right, \textit{Brown} indicates that this is not a right which extends to both heterosexuals and homosexuals.\textsuperscript{22} It is difficult to identify the reasons for confining the argument of privacy to husbands and wives only.\textsuperscript{23} The case of \textit{Brown} was taken to the European Court of Human Rights (ECtHR) where it became \textit{Laskey et al. v UK}. The ECtHR held that the CA’s decision to criminalise consensual sado-

\textsuperscript{17} Sexual Offences Act 1967.
\textsuperscript{19} W Wilson, \textit{Criminal Law: Doctrine and Theory} (2\textsuperscript{nd} edn, Pearson Education, Harlow 2003) 335.
\textsuperscript{20} \textit{Brown (Anthony Joseph)} [1994] 1 AC 212 (HL) (Lord Jauncey of Tullichettle).
\textsuperscript{21} \textit{Wilson (Alan Thomas)} [1996] 2 Cr App R 241 (CA).
\textsuperscript{23} A Ashworth, \textit{Principles of Criminal Law} (5\textsuperscript{th} edn, Oxford University Press, Oxford 2006) 323.
masochism violates Article 8, however went on to note that the state’s interference was justified as there is a need to regulate violence of this kind. More importantly however, the Strasburg Court was urged to acknowledge that English Law is prejudiced against homosexuals; however the Court answered that the HL considered only the seriousness of violence involved rather than the sexuality of those involved.\textsuperscript{24} Consequently, consensual homosexual conduct has been decriminalised which is an enormous step forward. However it should not stop there, the law should continue to be transformed in order to provide equality to homosexuals in all areas of law.

1.3. **Provocation under the Homicide Act 1957**

The unfair discrimination of homosexuals continues to remain extremely low on the judiciary radar; not only has there been a failure to draw a distinction between homosexuals and heterosexuals by the English courts, but the legislative protection afforded to homosexuals who kill their partners as a result of domestic abuse is restricted and practically non-existent. For years feminists and academics such as Mackie have strongly asserted that the law contains gender biases against women, especially surrounding provocation.\textsuperscript{25} In *Smith (Morgan)* Lord Hoffman recognised that provocation has deep seated logical and moral flaws. He further stated that the law embodies Kant’s dictum that, “from the crooked timber of humanity, nothing completely straight can be made.”\textsuperscript{26} While these gender biases are evident, one advocates that the biases against homosexuals are more evident, damaging and consequential. Were the internal elements of provocation successful in reaching its aims, the secondary consequences remain intolerable. Provocation will always encourage not only sexism and racism, but homophobia also.\textsuperscript{27}

\begin{footnotes}
\item[24] Laskey et al. *v United Kingdom* (1997) 24 EHRR 39 (ECtHR) [47].
\item[26] Smith (Morgan James) [2001] 1 AC 146 (HL) [Lord Hoffman].
\end{footnotes}
Before 2009, the law on provocation was governed by section 3 of the Homicide Act 1957. Provoked killings are largely considered less horrific than unprovoked killings, and provocation was accepted as just grounds for reducing murder to manslaughter. To satisfy this partial defence to murder, the accused must satisfy two main elements: the subjective requirement, and the objective requirement. The subjective limb demands that the defendant was provoked in such a way that would cause him to lose his self-control and kill. The objective limb requires that further to the subjective, the defendant must also have lost control in the same way that a reasonable man placed in the defendant’s circumstances would have done.

1.4. Loss of Control Under The Coroners and Justice Act 2009

A 2004 Law Commission report and the case of Holley recognised the problems associated with provocation. Lord Nicholls advocated that the state of the law was not satisfactory and that it was “flawed to an extent beyond reform by the courts.” The three main problems identified by the Law Commission were that: provocation had become too unconstrained as often the judge leaves the matter to the jury where conduct or words depended on are trivial; the loss of self-control was considered to be gender bias, and that slow-burn cases extended the law beyond its boundaries; and the objective requirement defined in Morgan (Smith) was too wide.

Following a Government consultation paper in 2008, a radical change was decreed upon provocation. Section 56 of the Coroners and Justice Act 2009 (CAJA 2009) abolished the old partial defence of provocation in what has been

28 Homicide Act 1957 s 3.
32 Ministry of Justice, ‘Partial defences to murder: loss of control and diminished responsibility; and infanticide: Implementation of Sections 52 and 54 to 57 of the Coroners and Justice Act 2009’ (Criminal aw Policy Unit, Circular 2010/13, 4 October 2010).
33 Coroners and Justice Act 2009 s 56.
described as “the biggest shake-up of murder laws for a generation.” However, the CAJA 2009 has only a passing resemblance to the original recommendations of the Law Commission. In place of the abolished provocation is a new partial defence to murder namely; loss of control. This is an enlargement of the former law as it covers loss of control arising from both anger or outrage (present in the old provocation) and fear.

Loss of control requires that the accused experiences a loss of self-control on account of a qualifying trigger. The loss of self-control needs not be sudden anymore; however again, revenge killings are not excused through this partial defence. Regarding the qualifying triggers recognised by the Act, s.55 postulates that these occur in 2 circumstances; through fear of serious violence towards oneself from the deceased or others, or was derivable from things “done or said” which “constituted circumstances of an extremely grave character” and “caused D to have a justifiable sense of being seriously wronged.” Therefore, as stated above, loss of self-control can arise from fear or anger, however these must not have been encouraged by the accused. Lastly, unlike the old law, sexual infidelity no longer plays a part in determining loss of control. An objective test is still in force, albeit widened. The new objective requirement specifies that “a person of D’s sex and age, with a normal degree of tolerance and self-restraint and in the circumstances of D” would have reacted similarly.

Chapter 2: The Subjective Requirement
The subjective requirement of provocation has significantly expanded in the new loss of control partial defence. Under the old law, the accused was required to successfully argue that they were provoked into killing, having demonstrated a loss of self-control. New law dictates that in arguing in favour of the loss of

---

34 C Hope, ‘Battered women who kill abusers can now avoid jail’ Daily Telegraph (London, 5 October 2010) 8.
37 Coroners and Justice Act 2009 s 54.
38 Ibid, s 55.
39 Ibid, s 55(4).
control defence, the accused must have experienced a loss of self-control as a result of a qualifying trigger. The subjective element has been enlarged to include a loss of self-control arising from fear, as well as anger and outrage which was stipulated under the old law. This chapter evaluates these changes and argues that the wider subjective test under the new law provides greater support to homosexuals who kill abusive partners following cumulative domestic violence.

2.1. Loss of self-control
The subjective requirement of the old law necessitates that the defendant was provoked into losing self-control and subsequently killed. In Acott Lord Steyn explained that provocation requires some form of human action, and cannot simply be a build-up of emotions. With regards to the loss of control experienced, a loss of temper alone is not enough. It must be that one is unable to suppress violent inclinations apparent under normal circumstances, which undermines the accused’s self-control. Duffy narrowed this subjective element to avoid the use of the defence for motivational, planned or revenge killings by including the constraint that the loss of control needs be sudden.

The word ‘sudden’ is not mentioned in section 3 of the Homicide Act 1957, and tremendously limits the scope of the partial defence. In the case of Ibrams Lawton LJ supported this view. He voiced that where there is a time gap between the final provoking act and the killing, the sudden and temporary loss of self-control requirement in Duffy is absent. The sudden requirement suggests that retaliations must occur promptly after the provocation, rather than asking only if the accused was in an uncontrolled state of mind. This additional element lies at the heart of the inability of provocation to develop from its heterosexual male stereotype and to encompass cumulative provocation. A prolonged period of maltreatment is oft associated with the domestic setting; a rising social problem

---

40 Acott [1997] 2 Cr App R 94 (HL) 102.
42 Duffy [1949] 1 All ER 932 (CA).
which requires addressing.\footnote{C Wells, ‘Provocation: The case for abolition’ in A Ashworth and B Mitchell, \textit{Rethinking English Homicide Law} (Oxford University Press, Oxford 2000) 89.} The law fails to recognise that domestic violence is a problem in homosexual relationships and that on-going abuse can lead someone to kill their perpetrator.

Lord Taylor, on the other hand, acknowledged in \textit{Ahluwalia} that the time lapse between the provocation and killing is pertinent, but not a determining factor in finding a loss of self-control. This of course is provided that when the killing took place, the accused’s provocation caused a “sudden and temporary loss of self-control.”\footnote{Ahluwalia (Kiranjit) (1993) 96 Cr App R 133 (CA) 139.} The condition for suddenness has been criticised for being unprecedented and for avoiding other ways of tackling revenge killings. It also favours defendants with short tempers and those with physical strength to act quickly, which in most cases is a male or dominant partner. This thwarts the use of the partial defence for those with a slow-burning temperament who allow the abuse to develop and amplify before attacking their perpetrator.\footnote{A Ashworth, \textit{Principles of Criminal Law} (6th edn, Oxford University Press, Oxford 2009) 253.} This is very unfortunate for those with tolerance, as when a partner kills another, there has nearly always been a history of violence.\footnote{C Wells and O Quick, \textit{Lacey, Wells and Quick: Reconstructing Criminal Law – Text and Materials} (Law in Context, 4th edn, Cambridge University Press, Cambridge 2010) 780.} Women are often associated with a slow-burn temperament; however this thesis asserts that this should be similarly recognised in homosexuals experiencing long-term abuse. Men in homosexual relationships may also endure systematic violence and kill their perpetrator following a final provoking act.

In \textit{AG’s Reference (No 24 of 2003)} Lord Mantell permitted the use of provocation by a man who had been domestically abused and tormented by his wife for years.\footnote{Attorney-General’s Reference (No 24 of 2003) [2003] EWCA Crim 2451 (CA).} While this case mainly deals with the issue of sentencing, it was accepted the systematic abuse was a factor to be considered. This case shows that the courts have allowed provocation to be used by men who have been domestically abused. It should be expected from this that the law will continue to allow
provocation to be used by men who have been the victim of domestic violence, especially those in homosexual relationships.

The new law seeks to eradicate the flaws of gender biases. The CAJA 2009 removes the sudden requirement established in Duffy,\(^{50}\) and further implements the new requirement of a qualifying trigger.\(^{51}\) This has also been beneficial to homosexuals. Now that the retaliation no longer needs to be sudden, abused partners, male or female, have more access to use the loss of control defence. However, there is a risk that the new law is being used improperly, even by abused partners. The Law Commission understand that there continues to be a vital problem where “a defendant has killed while basically in full possession of his or her senses, even if he or she is frightened.”\(^ {52}\) This may therefore take on the form of a revenge killing rather than killing following a loss of self-control. On the one hand it is beneficial as anyone who kills out of anger must have, at some point, lost their self-control. On the other hand, it can still be hard to argue that the accused had lost their self-control when there had been a delay in reacting.\(^ {53}\)

2.2. Qualifying Triggers

As Lord Hoffmann stated, victims of domestic abuse are the kinds of people which provocation should seek to protect, irrespective of a direct relationship between provocation and characteristic.\(^ {54}\) Under the new law there are two paths available to battered partners. First is the new fear trigger, which asserts that the loss of self-control can be as a result of fear of serious violence.\(^ {55}\) Second is through a new interpretation of how domestic violence influences loss of control through anger.\(^ {56}\) This requires that the defendant loss self-control following things said or done (or both) which either “constituted circumstances

\(^{50}\) Coroners and Justice Act 2009 s 54(2).
\(^{51}\) Ibid, s 54(1)(b).
\(^{54}\) Smith (Morgan James) [2001] 1 AC 146 (HL) 305-307 (Lord Hoffmann).
\(^{55}\) Coroners and Justice Act 2009 s 55(3).
of an extremely grave character,” 57 or caused the defendant to “have a justifiable sense of being seriously wronged.” 58

The fear trigger was first mentioned by Lord Hoffmann in Morgan (Smith) where he said that emotions “may cause [a] loss of self-control [and] are not confined to anger but may include fear and despair.” 59 However at the time, this was rejected. It was then presented by the Law Commission to deal with cases which did not satisfy self-defence, as often, when in fear of violence, retaliation is not proportionate. The new loss of control defence therefore represents defendants who may have overreacted or made a pre-emptive strike. 60 Including fear in the qualifying triggers is good for homosexuals subject to long-term abuse as they will be aware that they will continue to experience violence and will be forever fearful of the next attack, which can lead to a loss of self-control.

The introduction of anger as a qualifying trigger seeks to bring abused partners within the narrower perimeters of the old defence of provocation. Often, following Holley, the changes to the defendant’s personality caused by on-going domestic violence leads to feelings of helplessness. 61 Yet helplessness itself can be a factor which causes a loss of self-control and so may require the use of the loss of control defence instead. Evidence of regular abuse does not affect one’s capacity for self-control but the magnitude of the mistreatment endured. Therefore, systematic abuse establishes the gravity of the offence and the justifiable sense of being seriously wronged which would have been similarly experienced by the reasonable man with normal tolerance and self-restraint. This argument circumvents the need to use diminished responsibility and label the accused as mad. 62 Following the new law, defendants can depict themselves as regular people who have suffered long-term abuse and so acted out of a

57 Coroners and Justice Act 2009 s 55(4)(a).
58 Ibid, s 55(4)(b).
59 Smith (Morgan James) [2001] 1 AC 146 (HL) 168 (Lord Hoffmann).
justifiable sense of anger at their treatment.\textsuperscript{63} This shows the great benefit of the new law and how it has come to provide safeguards to those who kill following systematic domestic violence. This part of the defence is also readily available to defendants of any sexual orientation and on the face of it, appears to not discriminate. As anger is not a stereotypical emotion, homosexuals are able to argue that they lost their self-control due to the anger aroused from their abuse. The requirement that the defendant had a justifiable sense of being seriously wronged guides the jury to evaluate what is morally and politically acceptable.\textsuperscript{64} While this is an objective standard, the recognition of long-term abuse has a great impact on the defendant’s ability to show that they have been seriously wronged by their abuser. This greater emphasis on morals may be as a result of a cultural change and a higher acceptance of acts which, in the past, were considered to be immoral or unacceptable. The acceptance of homosexuality has been a great cultural change since 1957, and this new requirement directs the jury to keep in mind that everyone is different. In this way, the jury should be made aware of the defendant’s sexual orientation and the nature of the relationship in which he was a partner to.

\textbf{2.3. Battered Woman Syndrome}

Throughout the 1980s the courts struggled to adapt to the social reality of domestic violence, and so the concept of battered woman syndrome (BWS) emerged. BWS was developed by a psychologist who intended to “map the stages through which women subjected to long-term physical abuse within a relationship might pass.”\textsuperscript{65} The syndrome advances that as a result of recurring violence, a woman is likely to become “immobilised, passive and unable to act to escape the oppression of her situation.”\textsuperscript{66} Women subjected to systematic abuse will usually not react to the first few provoking incidences, but may eventually

\begin{footnotesize}
\textsuperscript{63} A McColgan, ‘In Defence of Battered Women who Kill’ (1993) 13(4) OJLS 508, 518.


\textsuperscript{65} C Wells and O Quick, \textit{Lacey, Wells and Quick: Reconstructing Criminal Law – Text and Materials} (Law in Context, 4\textsuperscript{th} edn, Cambridge University Press, Cambridge 2010) 781.

\textsuperscript{66} Ibid, 781-782.
\end{footnotesize}
snap after a further provoking incident, even if it was relatively minor.\textsuperscript{67} Many women were unable to successfully plead provocation as they could not satisfy the subjective and objective elements required for the partial defence. Therefore, many women would turn to BWS to support their claims of provocation.

\textit{Luc Thiet-Thuan} accepted that BWS is relevant to “the first and subjective inquiry,” but not the “second and objective inquiry.”\textsuperscript{68} Therefore, BWS can be helpful as it provides women with a medicalised specialist term that can be admitted into evidence to assess the subjective requirement of provocation. That is to say that evidence of BWS can help support a claim that a woman lost her self-control.\textsuperscript{69} However this has caused problems as the use of expert evidence makes the woman’s ‘syndrome’ appear unusual and so it may be better suited under a defence of diminished responsibility. Unfortunately, to admit evidence of BWS into a claim of diminished responsibility would place the user into an unstable typology. Further, this approach is internally inconsistent because BWS relies on a woman’s vulnerability whilst simultaneously explaining that she killed.\textsuperscript{70}

Despite the criticisms of BWS, it is nonetheless supports a woman’s defence of provocation. However, no such medical explanation exists for homosexuals who experience domestic abuse and act similarly to women. As stated above, there are various reasons why homosexuals may find it difficult to leave an abusive partner, and so may suffer at the hands of their perpetrator for a number of years.\textsuperscript{71} It is advocated that BWS should recognise a slow-burn temperament in all sufferers of long-term domestic abuse. Rather than only having regard to the sex of the victim, BWS should evaluate the nature of the relationship and recognise that it could be experienced by anyone, regardless of gender or sexual orientation. When women kill, they are usually responding to a high degree of masculine violence aimed towards them throughout their relationship with the

\textsuperscript{68} Luc Thiet-Thuan v R [1997] AC 131 (PC) (Hong Kong) 150.
\textsuperscript{69} Ahluwalia (Kiranjit) (1993) 96 Cr App R 133 (CA).
\textsuperscript{70} C Wells and O Quick, \textit{Lacey, Wells and Quick: Reconstructing Criminal Law – Text and Materials} (Law in Context, 4\textsuperscript{th} edn, Cambridge University Press, Cambridge 2010) 782.
\textsuperscript{71} C Renzetti and D Curran, \textit{Women, Men, and Society} (4\textsuperscript{th} edn, Allyn and Bacon, USA 1999) 185.
perpetrator.\textsuperscript{72} In the same way, as it has already been revealed, abuse from a dominant partner is just as common in homosexual relationships.\textsuperscript{73} BWS should therefore not solely be about the gender of the victim of abuse, but rather the nature of the relationship which has been endured. The law needs to recognise that the slow-burn temperament that BWS finds in women can be similarly found in men.

Nonetheless, even if men can use BWS it still has all of its flaws: It fails to reflect the true nature of abusive relationships; it is methodologically unsound; and it incorrectly pinpoints the pathology in the sufferer of disordered behaviour.\textsuperscript{74} Downs further identifies the syndrome society which, through “syndromising” patterns of behaviour, relocates the blame from the perpetrator to the victim. Downs argues that the perimeters of the syndrome are too narrow, and as a result, fails to protect all those who should benefit from it.\textsuperscript{75} Whilst Downs is arguing here that not all women are included, this thesis asserts that unfortunately, BWS does not recognise or help all of those experiencing long term abuse such as those in same-sex relationships. Nevertheless it is important to be reminded that the uncovering of domestic violence and the emergence of BWS demonstrate a development of awareness of gender.\textsuperscript{76}

Discussions regarding whether BWS will be effective under the new partial defence of loss of control is likely to become a highly contentious issue. As previously stated, it may be hard to show that, even with the removal of the suddenness requirement, one can lose their self-control when there has been a delay. However cumulative abuse can cause a loss of self-control even if the fatal incident was minor.\textsuperscript{77} Therefore, it can be argued that BWS is needed to explain
these reactions. Perhaps instead a similar medical explanation should emerge that recognises that submissive partners can experience a delayed response, regardless of sex or sexual orientation. The law should be prepared to recognise that what BWS acknowledges can also be found in men and in homosexual relationships. Conversely, it could also be argued that BWS should be removed altogether and instead the courts should have full regard to the nature of the relationship endured when having regard to whether the defendant lost their self-control.

2.4. The Homosexual Advance Defence

Unfortunately, the law on provocation was not so kind to homosexuals and contained direct discrimination. The homosexual advance defence (HAD), sometimes referred to as the Portsmouth defence, was a partial defence to murder under provocation. If successful, murder would be reduced to manslaughter where the defendant showed that “the victim made an unwanted homosexual advance so repugnant to the defendant that it provoked him into an excusable lethal attack.”78 Whilst elsewhere in the world the use of the HAD has been endorsed, the stance of the English courts is unclear.79

In the case of Green the Australian High Court remarked that male honour must receive a stamp of approval.80 However in England and Wales, there have been no cases which specifically deal with this issue. The few cases which discuss the HAD have been concerned with sentence rather than its suitability as a plea of provocation.81 However the courts in Moreby appeared willing to consider the HAD. In this case the deceased had approached Moreby and asked if he was gay. When the deceased was declined, he followed Moreby and grasped his crotch.82

In sentencing Moreby to four-and-a-half years for manslaughter, the court

---

81 McCarthy [1954] 2 QB 105 (CA).
82 Moreby (John Peter) (1994) 15 Cr App R (S) 53 (CA).
showed its readiness to accept the HAD notwithstanding the absence of proportionality.\(^{83}\)

With the subjective element in mind, it is hard to see how the defendant lost control as it could also be identified as prejudice and contributing in a hate crime. By approving the HAD the courts would be justifying behaviour which seeks to attack sexual orientation. This is an undesirable result. The new subjective element also fails to eradicate these problems as it could be argued that the homosexual advance by the victim caused the defendant to become angry and feel seriously wronged. Thus, a loss of self-control could be found. On this basis, both provocation and loss of control have failed to distinguish between excusable violence and blind prejudice.\(^{84}\)

Whereas discrimination on the grounds of racism or religious hatred is generally rejected, homophobia has not been fully erased. A trial judge may look to the current distinctions between homosexuals and heterosexuals given in Brown and Wilson and may note the lesser degree of protection afforded to homosexuals. This contrast might be used as reason to leave the jury to decide the appropriateness of such a defence rather than withdrawing it from the jury.\(^{85}\)

Taking into consideration that provocation needs not be proportionate, if the jury say it is fair, then the plea will be successful. While Duffy disputed this, commenting that provocation should encompass proportionality,\(^{86}\) this was overruled.\(^{87}\) Instead the approach in Scots law should be followed as it is fairer to homosexuals due to the necessary element of proportionality. Rejection of the relevance of proportionality allows for the defence to render the life of a homosexual inferior than that of a heterosexual man.\(^{88}\) Where the homosexual defence is permitted it insinuates that the courts are directly endorsing

\(^{84}\) C Wells, 'Provocation: The case for abolition' in A Ashworth and B Mitchell, Rethinking English Homicide Law (Oxford University Press, Oxford 2000) 95.
\(^{86}\) Duffy [1949] 1 All ER 932 (CA) (Lord Devlin).
\(^{87}\) Brown (Egbert Nathaniel) [1972] 2 QB 229 (CA).
favouritism and bigotry, and provocation is revealed to be even more wrongfully discriminating.

Chapter 3: The Objective Requirement

In its 2003 report, the Law Commission recognised that the objective requirement defined in Morgan (Smith) was too wide as it allowed “a defendant to rely on personal idiosyncrasies which make him or her more short tempered than other people.” However, it is inherently difficult to change the law of provocation and remove the biases without questioning the whole structure of homicide. Factors such as motive, emotion and personal mitigating circumstances require evaluation. Any changes made may not nullify discrimination, and homophobia could continue to occur where satisfactory changes to the law surrounding homicide are made. This chapter highlights the changes made and recognises that while discrimination continues to be an issue, the changes made have made significant progress in eliminating any bias.

The objective condition of provocation required that once the court was satisfied that the defendant was provoked to lose self-control, it must secondly decide whether the provocation was enough to make a reasonable man act in the same way as the accused. This can be subdivided into three categories for consideration: the gravity of the provocation; the adequacy of the level of self-control in response to the provoking act; and whether the provocation was grave enough to make a reasonable man act as the defendant did. The objective test apparent in the old law continues to be stipulated in the new law, therefore many of the problems to be addressed are still causing havoc. However, the test has been widened to include a reasonable man of the defendant’s age and sex, as well as having a normal degree of tolerance and self-restraint under the same circumstances experienced by the defendant. These new grounds of

92 Coroners and Justice Act 2009 s 54.
consideration for the objective test can be both beneficial and detrimental to homosexuals in a number of ways.

3.1. The Gravity of the Provoking Act
Under both the old law and the new law, the court must have regard to how grave the provocation was. In DPP v Camplin Lord Diplock held that the reasonable man should be someone of the defendant’s age and sex, whilst sharing the “accused’s characteristics as they think would affect the gravity of the provocation to him.”93 Norrie argues against Lord Diplock, suggesting that the issues of ‘provocability’ (one’s self-control) and ‘provocativeness’ (the gravity of provocation) are not inclined to be segregated so accurately.94 The CA developed the Camplin test in Newell where, following New Zealand Law,95 it was held that only permanent characteristics such as race and disability should be taken into account when considering the gravity of the provoking act.

Under the old law, for the characteristics to be relevant to the case, the provocation must have been aimed at it.96 This can be derived from the wording of the provision which states that in determining whether the provocation was enough to make the reasonable man act as the defendant did, the jury should “take into account everything both done and said according to the effect which, in their opinion, it would have on a reasonable man.”97 In other words, other than age and sex, only characteristics which were being taunted could be considered in the case.98 For example, in Morhall the reasonable man was held to be a glue-sniffer because the defendant had been mocked about his addiction.99 Therefore, the only way that a person’s homosexuality becomes a relevant issue is if the deceased taunted the accused about it.100 The implication of such a judgement is that these matters are left for the jury to consider with little

93 DPP v Camplin [1978] 2 All ER 201 (HL) (Lord Diplock).
96 Newell (Alan Lowrie Alastair) (1980) 71 Cr App R 331 (CA).
97 Homicide Act 1957 s 3.
99 Morhall (Alan Paul) [1995] 3 All ER 659 (HL).
guidance. Ashworth disapproves with this state of affairs. He expresses that “there ought to be a normative element that excludes attitudes and reactions inconsistence with the law or inconsistent with the notion of a tolerant, pluralist society that upholds the right to respect for private life without discrimination (Articles 8 and 14 of the Convention).”101 However, in the previous Morhall hearing, the CA accepted that homosexuality was an example of a characteristic which is in keeping with the reasonable man, and so should be left to the jury to consider.102 The implementation of such a statement has yet to be considered. Following previous standing on the criminal law’s view of homosexuality it seems unlikely to carry much weight.

While this position has not experienced any major change, the wording of the CAJA 2009 suggests a slight variation to the previous standpoint. The CAJA 2009 states that “the reference to “the circumstances of D” is a reference to all of D’s circumstances other than those whose only relevance to D’s conduct is that they bear on D’s general capacity for tolerance or self-restraint.”103 This suggests that there has been an expansion of the law. Whilst characteristics that only impact on the defendant’s capacity for self-control continue to be excluded, the Act does not impose any other restrictions on the consideration of any other characteristics. Therefore, it may be possible for the sexual orientation of the defendant to be included in an evaluation of the defendant’s circumstances.

3.2. Adequacy of Self-Control
The court must also consider how adequate the level of self-control was in response to the alleged provocation. Individual factors such as age and sex should be acknowledged here as these will provide the justification for the level of self-control exerted. Luc Thiet-Thuan made the apparent harsh decision that mental conditions or personality disorders which may affect one’s capacity for self-control should not be taken into account when regarding the adequacy of self-control.104 To the contrary, the majority in Smith (Morgan) were of the

102 Morhall (Alan Paul) [1993] 4 All ER 888 (CA) 893.
103 Coroners and Justice Act 2009 s 54(3).
104 Luc Thiet-Thuan v R [1997] AC 131 (PC) (Hong Kong).
opinion that juries should take account of all characteristics.\textsuperscript{105} This discrepancy was resolved in \textit{AG for Jersey v Holley} where a majority in the Privy Counsel preferred \textit{Luc Thiet-Thuan}'s decision.\textsuperscript{106} In his judgement, Lord Nicholls held that \textit{Morgan Smith} “involves a significant relaxation of the uniform, objective standard adopted by Parliament...the statute does not leave each jury free to set whatever standard they consider appropriate in the circumstances by which to judge whether the defendant’s conduct is ‘excusable.’”\textsuperscript{107} Nevertheless, a problem with following the decision in \textit{Holley} and \textit{Luc Thiet-Thuan} is that the capacity for loss of self-control is not only demonstrated by the accused’s capacity, but by external factors also.\textsuperscript{108} This combination is imperative for cases containing battered partners who are subject to systematic abuse where capacity for self-control is affected by their apprehension and knowledge of the perpetrators conduct. This constructs the defendant’s decisions and plan of action.\textsuperscript{109}

Many women who had been provoked into killing their perpetrators found it extremely difficult to satisfy either the subjective or objective elements of provocation. Provocation was both sexist and homophobic as the loss of self-control element favours paradigmatically heterosexual, male violence as explosive rage is typical of heterosexual males. Edwards saw provocation as “a validation of male experience, transforming men’s experience into an “objective” doctrine which passes for the “normative.””\textsuperscript{110} Unfortunately, as this area of the objective test remains the same under the new loss of control defence, these standards are still problematic and homophobic. Additionally, the objective requirement is culturally saturated as it invites the jury to assess whether the defendant can be excused of their crime, and when assessing reasonableness the jury will normally bring forward their cultural values. These values are usually white, male, heterosexual, middle class, and Christian. It is therefore extremely

\textsuperscript{105} \textit{Smith (Morgan James) [2001]} 1 AC 146 (CA).
\textsuperscript{106} \textit{Attorney-General for Jersey v Holley} [2005] 2 AC 580 (PC) (Jersey).
\textsuperscript{107} Ibid, [22].
\textsuperscript{108} \textit{Smith (Morgan James) [2001]} 1 AC 146 (CA) (Lord Millett).
difficult for homosexuals to fulfil these ideals and for the jury to accept that they would have acted in the same way as a ‘reasonable man.’¹¹¹ Power suggests that these problems are due to culture complexity and change. The focus of the law has shifted from the nature and source of the provocation, to the defendant’s state of mind when committing their offence.¹¹² This, according to Norrie, is “designed to avoid open and contentious moral and political issues.”¹¹³ Loss of control has therefore developed from justifying actions to excusing them based on heterosexual values.

Donald Nicolson discusses the inaccessibility of the law on provocation to battered women who killed their abusers. He explains that this emanated from two forms of gender discrimination namely; double-standards, and formal equality.¹¹⁴ Double-standard discrimination arises as women are expected to conform to standards not required of men. Formal equality recognises that requirements of behaviour standards rely on male conduct. Even though the law is not directly discriminatory towards women, it continues to be indirectly discriminating as it rewards women who conform to these labels, and punishes those who do not.¹¹⁵ It can be argued that in the same way, the law is also discriminate towards homosexuals as they are expected to conform to heterosexual standards, and are also punished for not matching the gender constructions.

The cases of Thornton and Ahluwalia illustrated further difficulties with the old law, which however, have been rectified by the new law. Both appeals alleged that the sudden requirement for loss of control obstructs the use of the defence where there has been a period of ‘cooling off’ and thus discriminated against those who act in such a way.¹¹⁶ Therefore, whilst differences between men, women, heterosexuals, and homosexuals were taken into account, sexual

---

¹¹² Ibid, 876.
¹¹⁵ Ibid, 185-186.
¹¹⁶ Thornton (No 2) (Sara Elizabeth) [1996] 2 Cr App R 108 (CA).
stereotypes were still in place and reinforced the oppression of homosexuals.\textsuperscript{117} Prosecution of women tended to become a trial of their character and their conformity with perceived standards of femininity, rather than a consideration of the facts of the case.\textsuperscript{118} It is believed that should this have gone before the courts concerning homosexuals, the treatment would have been the same – discriminate and unfavourable. Fortunately, The CAJA 2009 has removed the requirement that the loss of self-control be sudden.\textsuperscript{119} It has not stipulated a how much time must pass before a loss of self-control will no longer be recognised, but this is a definitive move away from enforcing stereotypes into the law, and provides those who experience cumulative provocation to be protected by the law.

A judge should instead have discretion in allowing some personal characteristics to be considered, a position which is closely linked to \textit{Camplin}. Similarly, the minority in \textit{Holley} were of the opinion that the jury should “take the defendant as they find him, ‘warts and all.’”\textsuperscript{120} The \textit{Camplin} test drew no distinction between characteristics affecting the gravity of the provocation and those affecting the general provocability of the accused.\textsuperscript{121} However as Lord Hoffmann in \textit{Morgan (Smith)} declared that characteristics such as jealousy or a bad-temper should not be considered, the new law in the CAJA 2009 is reflected in the decision. In the later case of \textit{Weller}, the CA held that the jury should be directed to take account of all matters, including “matters relating to the defendant, the kind of man he is and his mental state.”\textsuperscript{122} By evaluating the kind of man the defendant is, it will force the jury to have regard to his sexual orientation and his standing within his relationship. Evidence of long term abuse within a homosexual relationship may therefore be of concern to a jury when considering the adequacy of the self-control.

\textsuperscript{118} Ibid, 201.
\textsuperscript{119} Coroners and Justice Act 2009 s54(2).
\textsuperscript{120} Attorney-General for Jersey v Holley [2005] 2 AC 580 (PC) (Jersey) [18].
\textsuperscript{121} DPP v Camplin [1978] 2 All ER 201 (HL) 174 (Lord Diplock).
\textsuperscript{122} Weller (David Alan) [2003] EWCA Crim 815 (CA) [16].
These arguments are as applicable to the new law as they were to the old. Mackay et al agree that the law should return to the looser objective standard found in Morgan (Smith), arguing that personalities cannot be dissected under the strict objective test, so it is unjust to remove the defence from those unable to fulfil it. The law would therefore fail to make a sufficient concession to human frailty. Norrie recognises that a lot of power is left to the jury when determining what a reasonable person sharing the defendant’s characteristics would do and this makes more problems than it solves. It is argued that some characteristics bear more weight than others, especially sexual orientation, as this denotes the nature of the relationship endured.

Positive steps were taken in the case of HM Advocate v McKean where it was held that in the more enlightened days of sexual equality there is no reason why provocation should not be extended to homosexual couples. It was stipulated that such couples must live together and demonstrate their “ties of love, affection and faithfulness.” While is it good to see the courts taking practical steps in favour of homosexuals, this case is a Scottish case. In addition, Lord MacLean recognised that prior to this case, the law failed to extend to homosexuals. Whether the compassionate nature of Scots law can prevail in English law is a matter for time and growth.

3.3. The New Characteristics of the Reasonable Man

The definition of the ‘reasonable man’ has long caused complications and confusion in many areas of the law. The standard of the reasonable man is an anthropomorphic standard and excludes many types of people. By the twentieth century the reasonable man had developed into “the man of virtue”.

---

126 HM Advocate v McKean (Vikki Lynne) 1997 JC 32 (HCJ) 32.
127 Ibid, 33 (Lord MacLean).
which constitutes women, non-whites and homosexuals as the ‘other’ with whom it is difficult to associate.\textsuperscript{130} Although the reasonable man has expanded to include a person of the defendant’s age and sex, as well as having a normal degree of tolerance and self-restraint under the same circumstances experienced by the defendant,\textsuperscript{131} for homosexuals, this has both negative and positive consequences.

Age is an important contribution as it signifies maturity along with one’s capacity for self-control.\textsuperscript{132} While Norrie is unclear what contribution sex brings to the new law,\textsuperscript{133} it is suggested that sex plays a huge role in self-control. This is a positive step for battered wives as the female status is recognised, which may in turn accept any slow-burn temperament. However, for homosexual males, especially those in abusive relationships, this is not so accommodating. By labelling all males as the same, the law fails to distinguish between heterosexuals and homosexuals, as each may act differently. The law still expects homosexuals to be evaluated against a heterosexual norm. Gay males are more likely to experience domestic violence than heterosexual males, and are more likely to suffer on-going violence where unable to break away from such a relationship. Therefore, as stated above, submissive gay males are likely to experience a slow-burn temperament. By including sex in the objective test, homosexuals may find it difficult to plead loss of control where he kills his abusive perpetrator. It should be noted that the removal of the suddenness constraint may negate this from happening, however if the objective test fails to be fulfilled then the plea is not satisfied.

The introduction of the defendant’s tolerance and self-restraint under the same circumstances as the accused to the objective requirement is one which has proved most valuable to homosexuals. As made apparent by the Hart and Devlin debate discussed in the first chapter, a number of judges have shown little

\textsuperscript{130}Ibid, 95.
\textsuperscript{131}Coroners and Justice Act 2009 s 54.
tolerance towards homosexuals. This removes the judge’s ability to incorporate their own level of tolerance into the reasonable man. Through this, the law is able to complement the standards declared in the European Convention on Human Rights which prohibits discrimination\textsuperscript{134} and the Equality Act 2010 that advocates that there should be no discrimination on the grounds of sexual orientation.\textsuperscript{135} It is an effective ruling for homosexuals as it combats the difficulties found in BWS and the HAD. By including these standards, the defendant’s morals are reflected in the reasonable man. Therefore, the jury must assess a case on the basis of whether a reasonable man, with the tolerance and self-restraint of a homosexual who has endured abuse for a number of years, and has not been able to retreat from the relationship, would have acted in the same way. It also removes the use of loss of control for homophobic crimes such as the HAD.

While it may be perceived that the introduction of tolerance and self-restraint can only be a great step forward for homosexuals, in practice it is yet to be seen whether this will be a reality. The law is extremely ambiguous as to the definition of tolerance and self-restraint and leaves its interpretation solely for the jury. If tolerance and self-restraint are to merely be an element of self-control than they will not be as liberating as hoped. On the other hand they may be expected to depict their ordinary dictionary meaning. If this is so, then the increase of safeguards afforded to homosexual men who kill their abusive partners as a result of on-going abuse are becoming a reality. Collins English Dictionary defines tolerance as a “capacity to endure something, esp. pain or hardship,“\textsuperscript{136} and self-restraint as “restraint imposed by oneself on one’s own feelings, desires, etc.”\textsuperscript{137} Further to this, restraint is defined as, “the ability to control or moderate one’s impulses, passions, etc.”\textsuperscript{138} If the law follows these definitions, rather than intending tolerance and self-restraint to only mean self-control, then the enactment of the CAJA is of great benefit to homosexuals.

\textsuperscript{134} European Convention on Human Rights Article 14.
\textsuperscript{135} Equality Act 2010.
\textsuperscript{137} Ibid, 1397.
\textsuperscript{138} Ibid, 1312.
3.4. **Battered Woman Syndrome**

Following *Luc Thiet-Thuan*, BWS does not feature heavily in the objective test. This case held that BWS was not a relevant factor for consideration under the objective requirement.\(^{139}\) This is affirmed in the CAJA 2009 which specifies that, provided that it does not account for the defendant’s loss of self-control, factors of BWS are irrelevant to the objective test.\(^{140}\) While this appears damaging to both women and homosexuals, there are some positives to draw out from this decision. The objective element does include someone of normal degree of tolerance and self-restraint under the same circumstances experienced by the defendant. Therefore, the history of abuse experienced will be taken into account. As BWS demonstrates that having been subjected to years of abuse, killing one’s perpetrator is how some people react, a normal degree of tolerance and self-restraint can be deduced from this.

3.5. **The Homosexual Advance Defence**

The problems with such a defence have been noted in the previous chapter. Nevertheless, in the current social climate where the tolerance of homosexuals is high, it may be harder for a defendant using the HAD to successful fulfil the objective requirement of the reasonable man. Due to “cultural conditioning and differences in psychological characteristics across the genders” a trivial homosexual advance made towards men is viewed as a transgression of their masculinity.\(^{141}\) This may be why disproportionality in such cases has been tolerable. Whereas in the past it would have been reasonable for a man to be provoked by such an advance and to even kill as a result, in more socially accepting times it will become more and more difficult to say that these actions are reasonable, normal or rational.

As previously recognised, under the new law, the objective test was adjusted to include a person of the defendant’s age and sex, as well as having a normal degree of tolerance and self-restraint under the same circumstances experienced

\(^{139}\) *Luc Thiet-Thuan v R* [1997] AC 131 (PC) (Hong Kong).

\(^{140}\) Coroners and Justice Act 2009 s 54(3).

\(^{141}\) Ibid, 591.
by the defendant. By including a normal degree of tolerance, it is doubtful that the HAD would be a successful avenue under loss of control. The jury are now obligated to consider what a normal degree of tolerance is, and whether this was apparent in the defendant at the time of the offence. Although judges have been shown to be the least accepting of homosexuals, this will not be featured in the jury’s considerations. In times of social change and increased social acceptance homosexuality, which was previously classified as a psychiatric disorder, it has become morally, politically and legally acknowledged. Therefore, one argues that it will not be possible to successfully plead under the HAD as it is no longer regarded as a normal degree of tolerance to be discriminate against homosexuals.

Conclusion

It is common knowledge that domestic violence is growing social problem within heterosexual relationships, yet the same violence experienced in homosexual relationships is often overlooked by the English legal system. A study by Henderson revealed that among the lesbians and gay men identified, 29% of the male sample questioned had experienced physical, sexual or mental abuse in a homosexual relationship. The Homicide Act 1957 failed to acknowledge the importance of protecting gay men and discriminated against them through the traditional values which it upheld. The Wolfenden Report shows that while it was established that homosexuality should be decriminalised, it was still generally agreed that it was immoral. Whilst legislation such as the Domestic Violence, Crime and Victims Act 2004 and the Civil Partnership Act 2004 recognises the rights of homosexuals, the criminal law has drawn distinctions between homosexuals and heterosexuals. These deep-seated values are evident in the cases of Brown and Wilson. Here the courts held that due to public interest,

---

142 Coroners and Justice Act 2009 s 54.
private acts of homosexual sado-masochism should be deemed unlawful,\textsuperscript{146} whilst private acts of sexual violence between heterosexuals should not be restricted in the same way.\textsuperscript{147}

The loss of self-control required under the old subjective test was extremely biased as it established a heterosexual and male standard. Duffy narrowed this further by necessitating that the loss of self-control be sudden.\textsuperscript{148} This therefore excluded many abused partners from arguing a defence of provocation as there was often a period of ‘cooling off.’ The new law however stipulates that the retaliation no longer needs to be sudden.\textsuperscript{149} This has been a great improvement to the law and now protects many battered partners regardless of their sexual orientation. Homosexuals are no longer judged by the heterosexual standards of the old law in this respect. The new law also introduces new qualifying triggers of fear of harm and anger, which can both amount to a loss of self-control.\textsuperscript{150} Both of these are beneficial to homosexuals who have experienced long-term abuse as fear of harm and anger at such treatment can lead to a loss of self-control and thus justify a feeling of being seriously wronged.

Provocation was too convenient for men who lost their temper and acted out of rage, and too slow to protect battered partners who suffer from cumulative domestic abuse. Although treated unfavourably, provocation gave more recognition to women through BWS which gave a medical explanation for their actions to help support a claim of provocation. Unfortunately, no such explanations were afforded to battered men in gay relationships who were beaten for years by their partners. Following the introduction of the new partial defence of loss of control, it is suggested that BWS should not solely be a gendered defence, but should instead have regard to the nature of the relationship experienced. However it should be noted that in doing so, BWS places a syndrome on the head of those who use it. Perhaps instead BWS should

\textsuperscript{146} Brown (Anthony Joseph) [1994] 1 AC 212 (HL).
\textsuperscript{147} Wilson (Alan Thomas) [1996] 2 Cr App R 241 (CA).
\textsuperscript{148} Duffy [1949] 1 All ER 932 (CA).
\textsuperscript{149} Coroners and Justice Act 2009 s 54(2).
\textsuperscript{150} Ibid, s 55(3) and (4).
be removed and the courts given an obligation to recognise the nature of the relationship endured when considering if the defendant lost control and whether someone with the defendant’s tolerance and self-control would have done the same.

The HAD is an example of where provocation directly discriminated against homosexuals as it permitted the use of the defence where unwanted homosexual advances were made to the defendant.\textsuperscript{151} However the stance of the UK courts is unclear. It would be hard to show that such actions were the result of a loss of self-control and not just a prejudicial attack or a hate crime. However, the introduction of tolerance to the new loss of control defence removes the use of the HAD due to a higher tolerance of homosexuality by the public.

The objective test of provocation was also a heterosexual standard, as it sought to ask whether the reasonable man would have done the same as the defendant when put into their circumstances. The problem with this is that the reasonable man was often interpreted to be a white, male, heterosexual, and so it was very difficult to relate that to a homosexual. When considering the gravity of the provoking act\textit{ Camplin} held that all characteristics that would affect the gravity of the provocation should be considered here.\textsuperscript{152} The wording of the new law allows for all characteristics to be considered, even if they were not taunted.\textsuperscript{153} This is another positive development for homosexuals, as their sexual orientation can now be regarded.

When regarding the adequacy of the self-control,\textit{ Luc Thiet-Thuan} held that only age and sex are characteristics which can be considered.\textsuperscript{154} However, this is not satisfactory as external factors, such as the nature of the relationship play an enormous role in one’s ability to stay controlled.\textsuperscript{155} It is argued that all characteristics should be considered here. The removed of the sudden requirement by the new law does improve matters for homosexuals in relation

\textsuperscript{152} DPP v Camplin [1978] 2 All ER 201 (HL) (Lord Diplock).
\textsuperscript{153} Coroners and Justice Act 2009 s 54(3).
\textsuperscript{154} Luc Thiet-Thuan v R [1997] AC 131 (PC) (Hong Kong).
\textsuperscript{155} Smith (Morgan James) [2001] 1 AC 146 (CA) (Lord Millett).
to the adequacy of the self-control however as a loss of control can still be found even when there has been a ‘cooling off’ period. Furthermore, the new characteristics which are present in the reasonable man under loss of control have provided more safeguards to homosexuals. By including the tolerance and self-control of the defendant, sexual orientation is taken into account as is the nature of the relationship endured. However the addition of the sex characteristic to the defendant may not be as beneficial as this reinforces stereotypes that all males react in the same ways and does not account for sexual orientation.

There are two cases which demonstrate the way in which the courts and the legislators should move forward and are of great importance. First, the Scottish case of *HM Advocate v McKean*, where it was held that now that society is more tolerant of homosexuals, there is no reason why provocation should not be extended to homosexual couples.156 This should equally apply to loss of control. Secondly, the case of *AG’s Reference (No 24 of 2003)*, which permitted the use of provocation to a man who was abused by his partner. This shows the need for loss of control to be equally available to men who have suffered at the hands of domestic violence, regardless of sexual orientation.

However, even with these added protections for homosexuals, Miles recognises that it is impossible to successfully change the partial defences to murder without considering the wider framework. During the passage of the Bill for the CAJA 2009, a former Law Lord, Lord Lloyd pronounced the modifications to the partial defences as a “dog’s breakfast.” Instead of the Law Commissions proposals, Lord Lloyd suggested that partial defences as a whole should be abolished and instead replaced with an “extenuating circumstances” proviso that could be added by the jury in a murder case thus awarding the judge with full sentencing discretion. Whilst this amendment was met with widespread approval, it unfortunately failed when put to a vote at Third Reading.157 One asserts that Miles is correct in believing that any changes to the partial defences

---

to murder will be unsatisfactory unless the wider framework is amended and corrected. This thesis suggests that in order to rule out discrimination in the laws on homicide against homosexuals the wider framework must be changed. Only when the whole framework is changed will the partial defences be readily available to anyone, regardless of sexual orientation.

This thesis argues that Wells has not been proven wrong when she states that she does “not believe that there has been a true penetration of the implications of gender, let alone race and sexuality in the way we understand criminal law.”

Overall, there has been definite improvement for homosexuals, even if it was not in the minds of Parliament to do so. While the modifications in the new partial defence of loss of control go some way to meet the criticisms of provocation and deliver a more moral result, more could be done to protect homosexuals who kill their perpetrator. Ultimately, it remains to be seen how the new law will be received by the courts and how the changes will affect homosexuals.

LAUREN CROOK

---


‘Live and Let Live': The responsibility of women in the maternal-foetal conflict

ELIZABETH REDINGTON

INTRODUCTION

As described by Shelley Burtt, at the heart of the maternal-foetal conflict are rival claims about “an individual woman’s responsibility to the foetus she carries. Does a woman’s ‘duty of care’ to her child begin at birth? Once the foetus is viable? Once she has decided to carry the pregnancy to term?” The purpose of this paper is to examine the current position on the laws regulating the rights of the mother in response to the foetal rights movement, and to determine the appropriate way forward for the United Kingdom.

Issues in the field of prenatal care have come to the surface in light of advancing medical technologies, and a push for increased awareness and enforcement of human rights. There is no contest that the pregnant woman in the maternal-foetal conflict is afforded numerous rights under international and domestic statute, and through her right to personal autonomy. The conflict arises when considering rights for the unborn child. On one side of the debate are foetal rights campaigners who believe the foetus should be granted the same rights as independent persons. Conversely, there are those who believe women should have the liberty to make the decisions concerning treatment inflicted on their own bodies, irrespective of the impact it will have on the future child they are carrying. John Stuart Mill maintained strong views against "paternalistic encroachments on individual liberty," which will ultimately be interpreted should the coercing and forcing of prenatal treatment and caesarean sections be legalised. However, as noted by Jennifer Jackson, women voluntarily inflicted with the condition of pregnancy who “reject advice and persist in risky habits… not only compromise their own future liberty of action, but also that of their families.” In reality, in an attempt to exercise personal autonomy by refusing

---

1 Shelley Burtt, "Reproductive responsibilities: Rethinking the fetal rights debate" p. 181.
treatment, the pregnant mother often ends up divesting her future right to be able to do so as a result of the potential injury or death of her and her child.

The risks of carrying a child to term and enduring labour will almost always be greater than terminating a pregnancy. Annually, 500,000 women worldwide die in childbirth and 1 in 10,000 women in the developed world will die as a direct result of pregnancy. If the pregnant woman is willing to accept such risks, should it be said then, that with this decision she also accepts responsibility for the risks involved in undergoing invasive procedures to protect the foetus she ultimately wishes to deliver?

The purpose of this paper is to examine the current position on the laws regulating the rights of the mother in response to the foetal rights movement, and to determine the appropriate way forward for the United Kingdom.

**CURRENT LEGAL STANDING OF THE FOETUS**

Within the United Kingdom and the majority of the United States, the foetus is not granted legal status or the enjoyment of rights until the moment of birth. Although this view has been upheld by case law and relevant statutory provisions, in light of a shifting societal focus on the protection of human rights this stance is increasingly becoming challenged.

There is conflicting international law makes the question of when the foetus becomes legally recognized. In the United Kingdom, the accepted view of foetal status is set out in numerous cases. In the 1992 case of *de Martell v Merton and Sutton H.A* Phillips J stated that “the human being does not exist as a legal person until after birth. The foetus enjoys no legal personality and an unborn child lacks the status to be the subject of legal duty.” He clearly outlined the position of the UK courts at this time, and this was confirmed shortly after in *Attorney General’s Reference (No. 3 of 1994)* where the (now widely accepted) “born alive” rule was

---

5 *De Martell v Merton and Sutton H.A.* [1993] QB 204 [217].
constructed. The ruling stated that “violence towards a foetus which results in harm suffered after the baby has been born alive can give rise to criminal responsibility.” This effectively illustrates the United Kingdom’s position that until a baby has left the mother’s body, it is not owed protection under the criminal law.

Scientific developments in this area and, in particular, related to the generation of and research on human embryos, have extended further the need for appropriate law. In the recent case of Evans v United Kingdom, Wall J stated that “there is abundant domestic authority, binding on [the court], that a foetus, at whatever stage of its development, has no existence independent of its mother. If a foetus has no right to life under Article 2, it is difficult to see how an embryo can have such a right.”

Throughout this judgement, Wall J makes a sound argument for the denial of rights of the embryo, however, he suggests that the significantly developed foetus and embryo are two wholly different entities which ought to be viewed separately in the eyes of the law.

Ireland’s position on the legal standing of the foetus will be discussed in more depth later in this paper, but they adopt a significantly more liberal view of the rights and interests that should be afforded to the foetus.

While the majority of the United States adopts a similar view to the UK, the extent of the protection offered to the foetus varies significantly from state to state. In 35 out of the 50 states a foetus is legally recognized as an entity able to be harmed, passing laws classing the killing of an unborn child as homicide. The landmark case of Roe v Wade afforded women the right to legally abort their foetus pre viability, and made sure to note that a foetus “is not to be classified as a person under the Fourteenth Amendment, and it has no independent right to life.” It was also noted, however, that after the foetus has reached viability the state has an “important and legitimate interest in protecting the potentiality of

6 Evans v United Kingdom [2004] 1 FLR 67 (Evans HC), per Wall J at para. 175.
7 Kevin Dalton, “Refusal of Interventions to Protect the Life of the Viable Fetus – a Case-Based Transatlantic Overview”, p. 2.
human life”\textsuperscript{8} and it may then “proscribe abortion” unless the life or health of the mother is at risk.

**RELEVANT BACKGROUND AND LITIGATION**

**MOTHER V FOETUS**

There have been numerous arguments for the rights of the foetus to be upheld on the basis of Article 2 of the European Convention on Human Rights. Article 2 (1) states that “Everyone’s right to life shall be protected be law.” Unfortunately, there is no further elaboration on who is included in the term ‘everyone,’ and it has been left to domestic courts to decide its meaning. In the United Kingdom, this decision came through the case of \textit{Paton v United Kingdom}\textsuperscript{9}. The commission in \textit{Paton} stated that Article 2, “by its nature, concerns persons already born and cannot be applied to the foetus.” The foetus was also excluded because “unless some specific exclusion were introduced into the article, it would prohibit abortion in all cases, even where the mother’s life was in danger.”\textsuperscript{10} Here we can see a distinct conflict between the rights of the mother and the rights of the unborn. If it were possible to formulate a law in which the current position on abortion was upheld but a greater scope of rights was given to the foetus, a situation in which we find a viable foetus (no longer able to be aborted by right) and a mother in severe medical danger will surely arise. Granting the foetus protection under the ECHR would satisfy the arguments of many activists, but in doing so it places the rights of a potential life over the actual rights of a woman in danger of her life. How can we justify placing one set of rights over the other? If the decision is made to favour the life of the foetus, will it ever be seen as morally acceptable to condone the denial of a pregnant woman’s right to life in order to protect the rights of an unborn child who merely has the potential to be born?

An argument refuting the movement for the rights of the foetus is formed around the simple fact that until a foetus leaves the mother’s womb, it is not yet an

\begin{itemize}
\item \textsuperscript{8} Ibid, p. 2.
\item \textsuperscript{9} \textit{Paton v United Kingdom} [1981] 3 EHRR 408.
\item \textsuperscript{10} Sarah Christie, "Crimes Against the Foetus: the rights and wrongs of protecting the unborn” at p. 67.
\end{itemize}
independent human being. In the United Kingdom and United States alike, a strong emphasis is put on the protection and enforcement of human rights, and many have argued in favour of protecting the foetus under these same rights. However, if a foetus has not yet reached ‘independent’ personhood or attained legal status as a human being, it would seem disjointed to impart upon them the same set of rights that apply to such individuals.

As discussed above, the United States has largely maintained a similar view on the rights of the foetus to that of the United Kingdom. However, during the recent Bush presidency a bill was passed increasing the protection of the rights of the foetus. The Unborn Victims of Violence Act (otherwise known as Laic and Conner’s Law)\(^\text{11}\) was passed in 2004, and has been seen as a step in the direction of granting the foetus personhood, although it was specifically expressed that the foetus is not to be classed as a person under the Fourteenth Amendment. The Act arose after the murders of Laci Peterson and her 7 and a half month unborn son, legally recognizing a child \textit{in utero} as a victim if killed during the course of any one or more of over 60 federal crimes of violence against the mother. When passing the Act, President Bush stated, “Any time an expectant mother is a victim of violence, two lives are in the balance, each deserving protection, and each deserving justice. If the crime is murder and the unborn child’s life ends, justice demands a full accounting under the law”.\(^\text{12}\) Currently, as mentioned above, 35 of the 50 States have adopted the law that an unborn child can be legally seen as a victim of homicide.

THE PERSONHOOD ARGUMENT

It has been said that, biologically, there is no contest that from the moment of conception an embryo is part of the human species. Rather, what we need to focus on is the moral sense of humanity. There are various opinions as to the critical point at which personhood arises, all of which must be explored before


moving forward to explore the foetus’ eligibility to enjoy human rights. Strong cases have been made by numerous academics, each encompassing a different moral standpoint with varied legal and societal consequences. In an analysis by Michael Tooley, it was controversially argued that until an individual is able to appreciate self-consciousness, they cannot acquire personhood. He has put forth the proposition that until an individual has reached personhood no right to life can be acquired, essentially claiming that no infant, born or unborn, has a right to life.

Unfortunately for Tooley, any support for his argument was ephemeral, because he failed to acknowledge broad public opinion, and was unable to define his concept in law. Indeed, it proved presumptuous to assume that the majority of individuals in modern society would accept a concept that effectively stands to justify potential infanticide, based simply on a lack of self-consciousness, even if this could precisely be defined. For example, once an individual has initially established personhood, they undoubtedly enter situations in which they lose self-awareness. When entering a state of sleep, an activity that is essential to survival, we lose awareness of ourselves, exemplified through a lack of consciousness and the presence of dreams. Surely Tooley would not assert through lack of self-consciousness one fails to maintain personhood when asleep, or obtruded for any number of other reasons?

A variation on Tooley’s proposition is the view that potential to develop into a self-conscious being gives one personhood, developed by Devine in The Ethics of Homicide. Again, this assertion has failed to gain traction, largely because of issues of definition. Devine’s argument also brings about complex moral issues, explored by Carson Strong in a hypothetical situation set in a laboratory. He asks us to imagine whether it is possible to keep an embryo alive in this laboratory, giving it the complete potential to develop into a healthy, self-conscious being. A sudden fire breaks out and when attempting to save what you can, you notice a 10 year-old child lying on the floor. You are faced with the decision to save either

---

the embryo (with its life support equipment), or the 10 year-old child. If the embryo enjoys the same moral standing as the child, the decision would have equipoise. However, Carson Strong argues with clear moral authority that there is a large difference in the moral weight attached to each being, and on that basis the morally correct choice would be to save the 10 year-old child.14

A third definition of personhood is put forward by Sumner, and supported by other academics,15 asserting that it is the acquisition of sentience which gives rise to personhood.16 Sentience is described as possessing “the capacity for feeling or perceiving”.17 Sentience has been explained as a minimalist way of describing consciousness, which compels us to revert to the views of Michael Tooley. Tooley’s initial arguments are supported by 20th Century philosopher and Professor of Philosophy Peter Singer. In my view (above), the arguments backing this view are based on shaky foundations, and have not achieved any substantial merit. For example, the capacity for feeling may represent a simply physiological response, rather than any inherent awareness of surroundings or ability to react to those feelings. Consequently, this idea forces us to take into consideration any being which may encompass the ability to feel or perceive. This moves the discussion away from the point of human personhood and encourages debate about the moral interests of animals and plants, which is beyond the scope of this review of human personhood.

The idea that in order for personhood to arise one must be involved in a network of social relationships, as part of a community, is expressed clearly by Mary Warren.18 This has not, however, been widely accepted: Scott is critical of this argument in a commentary on the personhood argument. She points out the flaws in Warren’s argument, describing a need to look beyond the language of Warren’s argument as “morally embarrassing.” When viewed from Scott’s

14 Carson Strong “A Framework for Reproductive Ethics” P.23 For a more detailed account of the hypothetical situation, see pp 22-24.
16 Ibid 24.
17 Ibid 24.
perspective, it seems hard to argue with the fact that the human community justification of personhood is, as with the many other arguments, insufficient in providing an adequate account of one of the many arguments for when personhood arises. How then should the moral conception of personhood be viewed? When presented with a diverse range of arguments with the only real common theme being inadequacy, for the sake of the foetus and the pregnant woman should we revert to what has been agreed on and impede the denigration of the law? Given the inconsistency of the rights of the foetus from country to country, the answer to this question appears to be a resounding ‘no’.

THE ISSUE OF ABORTION
The right not to reproduce comes hand in hand with the right to reproduce. While approximately 3 million babies are conceived worldwide each year through in vitro fertilisation (IVF)\(^19\) by couples who are not able to conceive naturally, over 200,000 abortions are carried out per year in the UK alone.\(^20\) A woman’s decision to opt for abortion is permitted by the Abortion Act 1967, though the various arguments for this law to be changed will not be specifically addressed in this paper. There is, however, a need to explore some of the concepts used to create the law surrounding abortion to potentially be used to reform the law concerning the rights of the foetus.

In the United Kingdom and the majority of the United States (established in *Roe v Wade*),\(^21\) a pregnant woman is permitted to request an abortion to be carried out without reason up until the foetus has reached the gestational age of 24 weeks, the end of the 2\(^{nd}\) trimester. However, in the UK, most hospitals and clinics will not consider termination of pregnancy beyond 18 to 20 weeks unless there is a risk of physical or psychological damage to the mother. Nonetheless, per the Human Fertilisation and Embryology Act, termination can be carried out beyond


the 24 week mark at the discretion of the doctor, if it is necessary to save the life of the woman, where there is evidence of extreme foetal abnormality, or when there is a grave risk of physical or mental injury to the woman.\textsuperscript{22}

The United States position on abortion (which changed through \textit{Roe v Wade}) was not decided without widespread and ongoing public debate. Fundamentally, the law generated by \textit{Roe v Wade} and the majority of public opinion has prevailed because the rights of the pregnant woman remain central to its philosophical basis. This was illustrated by the first real opportunity to overturn the Supreme Court decision in \textit{Roe} in \textit{Planned Parenthood of Southeast Pennsylvania v Casey}.\textsuperscript{23} Despite extensive arguments regarding the rights of the foetus, and extensive lobbying from ‘pro-life’ groups, the Supreme Court elected to retain the current law and maintain a ‘pro choice’ attitude to abortion, with the mother’s rights being regarded as paramount.

In Ireland, the Eighth Amendment of the Constitution introduced a constitutional ban on abortion, effected on September 7\textsuperscript{th} 1983, which expressed an explicit ‘right to life of the unborn.’ This addition was not made without extensive scrutiny, but it provides the foetus with the freedom of undisturbed potential to develop to term.\textsuperscript{24} This has led Asim Sheikh to suggest that in Ireland, where there is a conflict between mother and foetus, maternal autonomy is non-existent or severely limited.\textsuperscript{25} This was established further through the case of \textit{Attorney General v X}. Although the majority of the Supreme Court held that in limited situations, abortion will be approved if it there was a ‘real and substantive risk to [the woman’s] life,’ the dissenting judgement of Hederman \textit{J} substantiates the emphasis placed on the rights afforded to the foetus in Ireland. At paragraph 72 of his judgement, he states that:

“The right of life is guaranteed to every life born or unborn. One cannot make distinctions between individual phases of the unborn life before

\textsuperscript{22} Human Fertilisation and Embryology Act 1990.
\textsuperscript{24} Exceptions to the anti-abortion law have now been granted in cases of rape, and where the mother’s life is in grave danger.
\textsuperscript{25} Asim A. Sheikh ”Medico-legal issues and patient autonomy – here yesterday, gone tomorrow?” p. 54.
birth, or between unborn and born life...The extinction of unborn life is not confined to the sphere of private life of the mother or family because the unborn life is an autonomous human being protected by the Constitution...Therefore, no recognition of a mother's right of self determination can be given priority over the protection of the unborn life. The creation of a new life, involving as it does pregnancy, birth and raising the child, necessarily involves some restriction of a mother's freedom”.  

It is clear therefore that Ireland is one of the most uncompromising advocates of foetal rights in the developed world.

**THE MOTHER’S DUTY AND THE SPECIAL RELATIONSHIP**

As discussed earlier, the case of *Paton v United Kingdom* was of great significance, not only for the clarification of Article 2 of the ECHR but for the development of arguments for and against the enforcement of foetal rights. Although the intention of *Paton* was to provide the United Kingdom with a definition and analysis of the meaning of Article 2, some of the reasoning that led to the definition is debatable. It is noted that abortion would have to be prohibited in all cases, “even where the mother's life is in danger”; but this reasoning is flawed as it does not take into consideration the possibility for a separate clause in the law. Though this statement is prefaced with “unless some specific exclusion were introduced into the Article...”, the phrasing implies this idea would be too complex to entertain. Numerous academics have, however, developed solutions to this that could potentially grant the foetus additional rights, and impose a duty on the pregnant woman to maintain a standard of care for the foetus.

In an exploration of the relevance of abortion law by Rosamund Scott, arguments considering the relationship between the current law on abortion and the relevance it has in the maternal-foetal conflict are established and developed. It is noted that John Robertson argued “once a pregnant woman decides not to

---

27 *Supra* note 9 (*Paton*).
28 For a further discussion see Rosamund Scott, *Rights, Duties and the Body*, Ch. 5.
have an abortion, she effectively waives the right to harm the foetus and could be said to be under a legal duty to take all reasonable steps to promote its well being.”

Established in *Roe v Wade*, the position in the United States stands that a woman has the right to terminate her foetus up until the point of viability, and after this point (the end of the second trimester) the pregnant woman essentially forgoes her right to abortion as long as her life and health are not in danger. In an in depth analysis of the moral obligation on pregnant women to undergo foetal surgery, Katherine Knopoff also calls on Robertson's proposition. However, she counters this by raising an important point, questioning the logic of drawing a line at viability. She posits that he could "push the line back to the point where the woman intends to bring the foetus to term, which could be much earlier than viability.”

She also rightly noted that it is almost impossible to assess the point of intent, but remarks that once intent has been established, there is nothing avertin a woman from changing her mind and deciding not to bring her foetus to term.

Although the main focus of *Attorney General’s Reference (No. 3 of 1994)* was to emphasize that the foetus does not establish legal status until after birth, the judgment has, in parts, provided a strong basis for the opposition. At paragraph 26 of the judgement, Lord Mustill stated that;

"There was, of course, an intimate bond between the foetus and the mother, created by the total dependence of the foetus on the protective physical environment furnished by the mother, and on the supply by the mother through the physical linkage between them of the nutriments, oxygen and other substances essential to foetal life and development. The emotional bond between the mother and her unborn child was also of a very special kind. But the relationship was one of bond, not of identity.

31 Ibid Knopoff at p. 509.
32 Ibid Knopoff at p. 509.
The mother and the foetus were two distinct organisms living symbiotically, not a single organism with two aspects. The mother's leg was part of the mother; the foetus was not."\(^{33}\)

It would be misguided to fail to recognise the profound bond created between mother and foetus during the gestational period in the womb. It is because of this bond one would imagine and expect the pregnant woman to treat the health and well being of the foetus as paramount, but in practice this is often not the case.

Though persuasive arguments have been put forward denying the reasonableness of imposing the extra duties on the expectant mother, it seems clear given to the unique nature of the relationship between mother and foetus this course of action would not be unreasonable. Once again the issue of termination of the pregnancy is raised when bringing about a potential duty to not harm the foetus, but this can be remedied simply. Shelley Burtt plainly refuted the prevention of a duty on the basis of the legality of abortion by stating:

"[b]ecause abortion remains, for at least some period of time, a morally acceptable option, discovering one is pregnant does not itself impose a duty of care to the foetus. Nor does a choice to continue the pregnancy. The ground of prenatal responsibility is rather the reasonable expectation (whether welcome or not) that the pregnancy will result in a live birth and a child to be cared for."\(^{34}\)

This view encompasses a range of issues. If we assume Burtt’s view was the accepted position on the maternal duty, it answers the questions of what course of action to take in the case of unwanted pregnancies, unintended pregnancies, and pregnancies resulting in a severely damaged or handicapped child. The right to an abortion is not being denied in Burtt’s view by any means. In the cases of pregnancy through rape, or a manufacturing defect of birth control, abortion is still an option and the potential for additional maternal duties never takes flight. By stating that the special relationship is not created solely through the fact of


\(^{34}\) Shelley Burtt, "Reproductive responsibilities: Rethinking the fetal rights debate" p 181.
conception, Burtt creates a clear dividing line between mothers who intend to carry to term, and those who simply find themselves inflicted with the condition of pregnancy.

General abortion law maintains the option of abortion after the 2nd trimester at the doctor’s discretion, albeit in very limited circumstances. For example, third trimester termination can be performed when there is a clear-cut risk of serious handicap. This does challenge the concept of the maternal duty to the unborn foetus, as discussed above. While the woman will have assumed a responsibility not to harm the foetus whilst it continues to develop inside her (inferred from the point of gestation), it would be fair to accept the concerns of the woman for the child once it has been born. At risk of assuming a lower quality of life for those children who are born disabled or severely handicapped or ill, the pregnant woman may legitimately believe it is in the best interest of the child to terminate the pregnancy. There is, however, a significant difference between positively choosing to prevent suffering, and negatively acting in ways to the detriment of the foetus (such as drinking alcohol, smoking, taking illegal drugs) and passively allowing the future child to die in refusing to allow doctors to perform life-saving treatment.

In a commentary by Katherine Knopoff on whether a pregnant woman can morally refuse foetal surgery, the issues of positive and negative duties imposed on the mother are discussed. She argues that in the deontological context, the pregnant woman would be required to refrain from, or punished for “smoking, drinking alcohol, taking drugs, or engaging in strenuous activity” as this could potentially harm the foetus. She also suggests that not taking prenatal vitamins or eating healthily could also be considered harmful to the foetus. The special relationship created between the woman and foetus upon conception could potentially justify an increased level of duty. In the law of tort, special duties arise when “one person is dependent on another who is in a position of dominance or power to aid.” Knopoff noted a significant difference in the

---

35 Knopoff, “...Fetal Surgery” p. 519.
36 Ibid at p. 523.
relationship between the woman and unborn child in that “no one else can possibly undergo a surgery to help the foetus but the woman.”\textsuperscript{37} Despite advances in medical technology, only she can provide the foetus the aid it requires, leaving a convincing case to lean towards the implementation of an actionable duty of the mother to her unborn child.

The argument supporting the imposition of a duty of the mother to the foetus once she has reached viability is sound, and is one of the strongest in trying to determine a middle ground which encompasses the potential legitimacy of the foetus, while aiming to protect the rights of the mother. It is apparent that granting the foetus the same rights that apply to living persons would introduce a plethora of concerns, primarily the impact on the current rights of the woman. Imposing a duty to the foetus upon a pregnant woman once she has made the decision to carry the future child to term does not impose any obligations in excess of those she should reasonably have assumed in choosing to become a parent. Bar the situations in which conception of the foetus occurred by no desire of her own, it is not irrational to expect the woman to provide a similar standard of care for the developed foetus as she would for a newborn.

Sadly, the prevalence of alcohol-dependent and ‘crack babies’ is ever increasing, and the United States has seen a significant rise in cases questioning the liability of the mother who failed to provide sufficient care for the foetus \textit{in utero}. Few people would stand by while a mother fed her newborn baby alcohol or injected its young veins with heroin; so the argument stands, why should it be acceptable for her to do so while the baby is developing inside her, when the effects are just as harmful, if not more distressing? While the woman may not have intended to harm her foetus, she “engaged in the voluntary act of sexual intercourse that caused the foetus to exist: but for that act, the foetus would not be suffering”\textsuperscript{38}.

\textsuperscript{37} Ibid at p. 534.
\textsuperscript{38} \textit{Supra} Knopoff, note 30 at p. 522.
TRUST IN THE DOCTOR-PATIENT RELATIONSHIP AND THE CAESAREAN SECTION

Assuming that the pregnant woman and the foetus are two separate entities with their own legal interests, the relationship between medical professionals and the expectant mother becomes a complex one. An additional standard of care arises upon a doctor tending to a pregnant mother, and due to the sensitive nature of the process of pregnancy the mother-to-be must be able to place a significantly increased amount of trust in the doctor to act in the best interests of her and her child.

Jennifer Jackson recounts the opening remarks of the British Medical Association’s description of the doctor patient relationship in *Medical Ethics Today*\(^{39}\) as:

> “The relationship between patient and doctor is based on the concept of partnership and collaborative effort. Ideally, decisions are made through frank discussion, in which the doctor’s clinical expertise and the patient’s individual needs and preferences are shared to select the best treatment”.\(^{40}\)

In theory this should provide the patient with the knowledge to make the best decision. However, this choice is not always an easy one to make, and the patient does not always make the ‘right’ or best decision. In the case where the pregnant woman is the patient in question, she is not only making choices that will affect her own body. In a commentary from Dickens and Cook, it has been suggested that “it is legitimate and prudent to extend healthcare professionals’ dedication to pregnant patients to their unborn children.”\(^{41}\) In this respect an obstetrician would be expected to weigh the interests of the foetus and pregnant woman equally, but what must the course of action be when there is clear cause for intervention, and one of the parties’ rights will have to be denied?

In the case of *Dobson (Litigation Guardian of) v. Dobson*, the majority adopted the view that “the best course of action is to allow the duty of a mother to her foetus

---


\(^{40}\) British Medical Association.

to remain a moral obligation which, for the vast majority of women, is already
freely recognised and respected without compulsion by law." The court must question the course of action to be taken when it is apparent that the
mother is not acting in the best interests of the foetus. We must ask when, if ever,
is it appropriate for medical intervention to override the legitimate wishes of the
mother? Although there is potential for this to be justified in certain cases, such a
law would only lead to inevitable challenge in the courts, but also significantly
undermining the trust that pregnant women should be able to put in their
doctors. It would be irresponsible to put forward arguments in favour of the
rights of the foetus while denying the vast impact it could potentially have on the
pregnant woman involved. Although medical and technical advances have
allowed for the foetus to become "imageable, visible, tangible and operable in
utero" after being previously inaccessible, there is no means for the foetus to
be treated without directly interfering with the mother's body. This has been a
large cause for concern within the courts as it is almost impossible to ensure that
a pregnant woman will not be unfairly persuaded to comply with the opinion of
medical professionals over her own intentions; and in the event she chooses to
uphold her decision, that a court order will not be sought to override her
decisions.

The UK and USA alike have seen numerous cases brought to the courts
surrounding forced and coerced Caesarean section procedures. The cases in the
United Kingdom began when the courts saw the case of Re S (now widely viewed
and wrongly decided). Sir Stephen Brown P decided to grant an order for an
emergency Caesarean section to be performed on a fully competent woman who
was refusing on religious grounds. In Re MB the doctor’s opinion once again
dictated the course of treatment to be taken when a woman with a severe phobia
of needles refused a Caesarean section that would have saved the life of her
foetus. Although the Court of Appeal proclaimed that the law protects a pregnant

42 Ian R Kerr “Pregnant women and the ‘Born Alive’ rule in Canada” p. 717.
43 Supra Dickens and Cook, note 41 at p. 85.
44 Ibid, p. 87.
woman’s right to autonomy and bodily integrity, as Rosamund Scott has pointed out it seems “irrationality can be seen as evidence of incompetence.”

The tone of decision-making changed, however, when the courts heard the case of St George’s Healthcare NHS Trust v S. The courts held that in the case of a competent patient, a Caesarean could not be carried out without the woman’s explicit consent, even where her “decision to refuse treatment (might) appear morally repugnant.”

Often relying on Roe, the United States legislature has heard repeated actions of forced Caesarean proceedings reaching all the way to the Supreme Court. The 1981 case of Jefferson presented a woman at 39 weeks gestation with placenta praevia (a condition that precludes safe vaginal delivery of the foetus) who refused to undergo a Caesarean section on religious grounds. The hospital applied for an order to carry out the procedure, and the Supreme Court of Georgia decided in the hospital’s favour, stating that a temporary denial of maternal autonomy was “outweighed by the duty of the State to protect a living, unborn being from meeting his or her death before being given the opportunity to live.” In 1986 the District of Colombia saw a case in which the mother was forced to undergo an emergency Caesarean section due to inconclusive suspicion of infection in Re Madyun.

As was seen in the United Kingdom, the tone of decision-making seemed to change in the United States with the case of Re Baby Boy Doe. Similar to the UK case of St George’s Trust, it was held that “even though a pregnant woman’s

---

46 Rosamund Scott, Rights, Duties and the Body p. 174.
48 A complication in which the placenta is attached to the uterine wall close to or covering the cervix which can lead to ante partum hemorrhaging which poses a great risk to the mother, and can potentially be fatal. http://www.uabhealth.org/15407/
49 Kevin Dalton, “Refusal of Interventions to Protect the Life of the Viable Fetus – a Case-Based Transatlantic Overview” p. 4.
50 Jefferson v Griffin Spalding County Hospital Authority 274 SE 2d 457 (Ga 1981).
51 Re Madyun 114 Daily Wash L Rptr 2233 (DC Super Ct July 26 1986).
foetus may be harmed, she may refuse a Caesarean section during her pregnancy if she is competent to make the decision”.\(^{52}\)

This leads to the tragic case of *Re A.C.* While the crusade for the establishment and protection of foetal rights is often a genuine and noble one, one cannot avoid the reality of the devastating impact prenatal intervention has had in the past. Angela Carder was a 27 year-old woman in her 25\(^{th}\) week of pregnancy when it was discovered she was suffering from a terminal form of cancer. She was informed that she had an inoperable tumour in her right lung, and made the decision to move into palliative care. By doing so she hoped to extend her life in order to reach the 28-week mark of pregnancy to allow her foetus a better chance of survival should intervention become necessary. Subsequently, Carder’s condition began to deteriorate rapidly, and she was predicted a life expectancy of 24-48 hours. Against the advice of their own obstetricians, and against Carder’s wishes, the hospital sought a court order to perform a Caesarean immediately in order to save the life of the foetus.

Despite making it clear she did not wish to go through with the procedure,\(^{53}\) it was authorised by the court and the child was removed, but died shortly after delivery. The mother survived her operation long enough to become aware of the fact that she had lost her child, but died 48 hours after the procedure.

It is because of cases such as these that any legal advances for the acknowledgement of rights for the foetus must be considered in great detail. It is likely that this will not be the last time a mother is forced to undergo treatment to which she does not consent for the sake of the foetus. There has been a great deal of commentary on this topic from activists and scholars alike, and in her book Sheena Meredith poses some important questions:

“is any woman to be so sacrificed – or only a dying one? How soon must she be expected to die to make this acceptable – a day, a month, a year? If

---


\(^{53}\) She was unable to speak, but clearly mouthed the words “I don’t want it done”, as per Dr Weingold 573 A 2d 1241 (DC App 1990).
a court was prepared to see Carder’s life foreshortened for the sake of her foetus, what about a severely disabled woman with a low quality of life? How disabled? How low?"\(^{54}\)

There is a clear tension. On the one hand an expectant mother would unquestionably want to be confident that her personal needs would be met while in hospital. On the other hand, the very basis on which she seeks medical care is the baby she expects to deliver, so one would presume she anticipates the same level of care for her future child. From the above cases it is clear that the decision making in both the United Kingdom and the United States is far from perfect, and the causes for pregnant women to lose trust in their doctors become apparent. It is essential to note that in the cases of *Jefferson, Re Baby Boy Doe*, and *Re Madyun* that the mothers all successfully delivered their children either vaginally, or without medical complications, contrary to the predictions of the medical professionals. Furthermore, for doctors to be able to do their jobs properly, their patients must place complete trust in their competence; but in turn, for patients to receive proper care, they must be willing to accept their doctor’s recommendations.

This tension becomes manifest when a decision has to be made that results in a conflict between patient and doctor. When a potential life is in the grips of being lost it is easy to lose sight of the rights the pregnant woman. As seen in *St George’s Healthcare NHS Trust v S*, the decision to refuse treatment is ultimately the woman’s, even if that decision will result in fatality for both her and her unborn child. As Jonathan Herring has written, “The right of autonomy is greatly treasured in our society but sometimes it calls for the highest of sacrifices.”\(^{55}\)

**CONCLUSION**

The central issues raised in the maternal-foetal conflict stem not only from conflicting societal opinions, but also from a legal shortfall in the clarification and protection of the rights of the mother and unborn child. We must ask what


\(^{55}\) *Supra* Herring, note 45 at p. 441.
changes so substantially in a few moments to merit a complete reversal of the way the life of the child is viewed? It is the position of this paper that although in law the foetus has no material claim over rights to life, once it has passed a certain gestational age the life and well-being of the future child should be brought to the forefront when making medical decisions. Through an analysis of the current legislation in the United Kingdom, United States and Ireland, and an in-depth study of pertinent case law, it is apparent there are inconsistencies which must be remedied. While the United States are ever increasing promoters of the rights of the foetus, the current position in the United Kingdom is to protect maternal autonomy and allow the rights of the pregnant woman to trump those of the unborn.

It is submitted that a pregnant woman’s omission to provide reasonable treatment for the foetus is morally reprehensible and should be reflected as such in the law. The decision to carry a foetus to term is one that imposes heavy burdens on the mother-to-be. As the objective of pregnancy is the successful delivery of a healthy child, it is not unreasonable to impart on the pregnant woman a set of duties to maintain a reasonable standard of care. In the emergency situation where a Caesarean section becomes a necessity, the withdrawal of the rights of the pregnant woman in order to protect the foetus should not become the standard. However, it is submitted that where the lives of both mother and child are in grave danger, medical intervention may be in the best interests of both parties.

In the fight to advocate for an otherwise voiceless child, we must not get caught up in the emotional whirlwind and neglect to consider the rights of the mother. One cannot deny the material human and state granted rights of the pregnant woman, and although it would prove beneficial to impose certain additional duties on her, to completely deprive her of her autonomy would undermine the very basis for debate. Nevertheless, in light of advancing technology allowing intervention to become increasingly less intrusive, the interests of the unborn child must be taken into account.
As Knopoff remarks, “the time to consider and resolve the apparent conflicts... is before any rights get trampled.” 56 Reviewing the progression of the foetal rights movement, the coercion and enforcement of Caesarean sections and the ongoing abortion debate, it is clear that until a level of consistency is established and maintained, a breach of rights is inevitable.

ELIZABETH REDINGTON

56 Knopoff, “...Fetal Surgery” p. 540.
**Mental illness and diversion from the criminal justice system**

CELYN COOPER

The practice of Diversion, as an outcome and process, has stood in political flux over the last four decades and “is a matter that drifts in and out of political consciousness”.1 Diversion, in its broadest interpretation, is “a process whereby people are assessed and their needs identified as early as possible in the offender pathway (including prevention and intervention), thus informing subsequent decisions about where an individual is best placed to receive treatment”.2

Following reports of high incidences of suicide (and suicide attempts) in British prisons, and the increasing awareness of the over-representation of offenders in the criminal justice system (CJS) suffering from poor mental health, legitimate concerns manifested into a pressing call for policy change in the mid-1970s. Even now, relatively recent figures published by the Prison Reform Trust3 recorded a total of 23,026 incidents of self-harm within the prison population. With a combined population of around 82,000, that total figure accounts for one quarter of the population itself.

The recent recommendation for diversion, by virtue of Lord Bradley states that “there are more people with mental health problems in prison than ever...custody can exacerbate mental ill-health, heighten vulnerability and increase the risk of self-harm and suicide. The failure of mainstream agencies to deal with these [concerns]...means that the Prison and Probation Service are in many cases being asked to put right a lifetime of service failure”.4 The diversion

---

3 The Prison Reform Trust 2006.
of mentally disordered offenders (MDOs), from the CJS to the mental health system (MHS) is firmly back on the agenda.

However, this pressing call for change is not a “recent phenomenon”.\(^5\) From the 18\(^{th}\) century it was thought that “no care is taken of [mentally disordered offenders], although it is probable that by medicines, and proper regiment, some of them might be restored to their sense, and usefulness in life”.\(^6\) The initial developments of diversion stem from the Butler Report in 1975, which led to the systematic development of Medium Secure Units in “in-patient provisions”.\(^7\) The report subsequently initiated the “first organized scheme for court diversion” in 1989 which led to their “fairly rapid, albeit piecemeal expansion” in the 1990s. Consequently, the Home Office Circular 66/90\(^8\) provided guidance on the care and treatment of such offenders “from health and social services authorities, rather than within the criminal justice system”.\(^9\) In identifying the concept of diversion as one of the essential methods in addressing the mental health needs of offenders, a “nationwide provision of properly resourced court assessment and diversion schemes”\(^10\) was called for by virtue of the Reed Report in 1992.

With opportunities of diversion arising at all pre-sentence, sentencing and remand prison stages, it was stated as imperative for all organisations involved “at the interface [of] criminal justice and mental health”\(^11\) to co-operate and improve communications respectively. Crucially, the most important implementation was that of “the incorporation of prison health services within

\(^8\) Home Office Circular 66/90 ”Provision for Mentally Disorder Offenders”.
the NHS” which sought to rectify “the high level of psychiatric morbidity within the prison population”.

The (former) Health Secretary Andrew Lansley, and the Justice Secretary Ken Clarke have publicly announced a “£5m scheme to divert mentally ill offenders from prison”, “in creating up to a further 40 diversion sites” across the UK. However, the focus in reality should shine on the existing schemes which “fail the national aim to deliver a coherent experience of justice for individuals passing through the criminal justice system”. As the Sainsbury Centre for Mental Health duly noted, “existing diversion arrangements are not achieving their potential… [with] only 1 in 5 people benefiting from the opportunity”. The difficulties facing the Government in implementing this goal is self-evident from the vogue of diversion in political rhetoric over the past 30 years.

The Bradley Report, published in April 2009, comes at a time when “political change and financial deficits” plague every aspect of our lives. As we stand, the current prison population stands at over 85,000 and with only around 2,000 spaces left, the UK penal system is in dire need of reformation. Born was the “rehabilitation revolution” – with the cost of keeping a single prisoner incarcerated for a year estimated at £45,000, and “the estimated financial cost of re-offending...around £11bn”, prison is damned as a “colossal failure”.

---

As a Government commissioned report, Lord Bradley found that “one of the main problems with previous policy development has been the piecemeal approach that it has taken: government department, agencies and organisations working independently of one another, developing policies and practice in isolation, addressing one problem or one part of the system at a time”\(^{21}\). Combined with the Sainsbury Centre for Mental Health’s Report\(^{22}\), which claimed that “an increased investment of £10m [to its allocated £20m] would translate to a £700m saving per year in trial, unnecessary imprisonment and other costs”\(^{23}\), diversion has not only become the main instigating force for treating the increased number of people with mental health problems in the CJS, but it has become the focal financial point in criminal justice policy. Whether this financial gain translates into “the combination of diversion-oriented sentencing policy, flexible mental health law, and an NHS [which could] create [a] fertile condition in which to create good services and effective diversion mechanisms for mentally disordered offenders”\(^{24}\) is certainly open to debate.

The majority of legal provisions which govern the treatment of MDOs are contained within the Mental Health Act 1983 (as amended by the Mental Health Act 2007). Some MDOs “cannot be treated voluntarily [thus, they] must be dealt with under the compulsory powers provided [in the Acts]”\(^{25}\). As these powers do not extend and apply within “prisons, where any treatment without consent must be justified under the common law...Part 3 of the MHA 1983 contains the provisions relating to the detention and treatment in hospital of patients with criminal proceedings or under sentence”\(^{26}\).


Essentially, different provisions may apply at different points of the offenders contact within the CJS, and thus, for the purposes of this essay, I will primarily concentrate on the three areas “in which divisionary effort is generally concentrated: the police station, the court and the remand prison”.27

Sections 135 and 136 of the MHA 2007 confers powers upon the police to remove to a “place of safety” an individual who may reasonably be believed to be suffering from a mental disorder from a public place (under Section 136), or by the issue of a warrant from private premises (under Section 135).

Section 135(1) specifically addresses the procedures which must be satisfied in issuing a warrant from a Magistrate to enter private premises to remove an individual to a place of safety with intent to make an application under Part II of the MHA (Section 2 admission for assessment). Based on information that has been “laid on oath by an Approved Mental Health Professional...reasonable cause [must be stated] to suspect that a person believed to be suffering from mental disorder has been or is being ill-treated, neglected or kept otherwise than under proper control”.28 The individual, in this sense, need not be under any suspicion of committing an offence, merely that there is reasonably cause to believe that he is suffering from a mental disorder.

In exercising and executing the powers of entry granted under Section 135, the police officer in question must be accompanied by an AMHP and Doctor, and may, as stipulated in the Code of Practice, “carry out [the] preliminary assessment in the premises themselves”.29 If it is felt that further provision or assessment is required, he may be removed and detained to a place of safety for no longer than 72 hours.

29 10.4 Code of Practice on the Mental Health Act 1983.
Because this Section may be applied in the event that the individual may not be suspected of committing a crime, it is difficult to assess what effect or difficulty it may pose in Government attempts to initiate a national policy for Diversion from the CJS. One may suggest that it could be a form of criminalizing sufferers of mental health, even when it is “far more likely that [mental health sufferers] will be the victims of violence themselves than an average member of the public”. Is it an example of Government policy shifting from “a culture of welfare to a culture of control” in response to sensationalist media reports that mental health sufferers pose a risk to society? Certainly, the wording of the Section serves as an explicit illustration that such persons should be “under proper control”. Here the Bradley Report suggested that all agencies involved in Section 135 and 136 place of safety orders should work together in developing “an agreed protocol on their use...immediately commences to identify suitable local mental health facilities as the place of safety”. These are all anachronistic recommendations, repeated in vain if anything, where local mental health facilities are sparse to say the least. They are orders which are “prime example[s] of why the police and health services need to work so closely together”, but resources, time and will have shown to be limited and constrained.

Nevertheless, the presence of a doctor, as well as an AMHP serves to satisfy doubts that the Section will not serve effectively in at least identifying a mental illness which is in immediate need of treatment; precisely what the underlying force of diversion epitomises.

Section 136 provides similar powers to those contained in Section 135, but is applicable where the person is found to be in “a place to which the public have access”. Here, the officer “has the power...to take them to a psychiatric hospital

---

where they may be detained for up to 72 hours to allow them to be assessed for detention".33

This use of this section has significantly increased since the MHA’s enactment in 2007 – over 10,000 times a year. While it may, however, provide safety to the prospective patient and the public, it is not without its problems. There are issues in relation to its effectiveness, the availability of resources needed to satisfy its requirements and purpose, as well as controversial claims of its abuse “in cases of members of minority ethnic communities”.34

Statistics on Section 135/136 place of safety orders have tripled over the last 3 years with recent research illustrating an “increase of 16% to 8,800 from 7,500 in 2007/08”.35 However, 6,200 (78%) of orders did not result in the individual being formally admitted, and from the guide notes of the statistics provided these figures are only for such orders where the place of safety was the hospital. It casts doubt over the role of the AMHP in determining whether the individual should be detained for assessment. While it shows a “greater use of police powers to remove a person who appears to be mentally disordered”, how will they fare when greater authority is bestowed upon them to divert where they clearly find it difficult to make relatively accurate preliminary assessments now? There will always be a “better safe than sorry” argument here, but to the extent that these figures show suggests that more training must be given to those on the front line of diversion, if the Government’s implementations are to succeed.

Furthermore, a recent report by the Independent Police Complaints Commission found that individuals suspected of suffering from mental illness “are being held in police cells almost twice as often as health-based settings as “designated

places of safety” while they await medical assessments”. This is despite stated practice by the Home Office which provides guidance to the effect that “…police station[s] should be used as a place of safety only on an exceptional basis”. A prison cell is certainly “not designed to house people in need of urgent therapeutic care” after warranting authority to make an assessment in the first place: it defies both sections’ purpose.

“When mentally disordered persons are arrested and taken to police...custody, sufficient resources are, in theory available to ensure assessment and transfer to health care facilities where appropriate”. The Department of Health in 2006 issued £100m to facilitate the need to “expand the number of appropriate places of safety...and to improve psychiatric care units” to support these aims. However, the funding, as Mind’s policy director Sophie Corlett noted, “did not extend to cover staffing costs and in some areas places of safety were standing empty. A much needed resource is going to waste”. Ken Clarke proposed to transfer £5m to fulfil the national application of diversion – 5% of the figure already provided by the DoH. It is simply unfeasible to believe that a coherent, efficient and effective system of diversion can be created from that sum of money. There is a limit to what the CJS and the MHS can achieve in this sense.

It is the custody sergeant on call at the station who makes the constructive choice to inform the forensic medical examiner to assess the prisoner if he suspects that he is suffering from a mental disorder, and the forensic medical examiner in turn who “may advise the custody sergeant to contact an AMHP to

arrange an assessment under the MHA”. To all intents and purposes, this system of screening seems infallible. With an individual representing each agency involved, it’s harder to envisage where a prisoner would “slip through the net” – go through the system without detection of a legitimate disorder and an immediate need for treatment. That said, “in practice, this system [has] rarely worked”. David James suggests that because the custody sergeant is primarily concerned “with legal matters about the person’s detention and to dispose of those arriving in custody as quickly as possible”, his knowledge of mental disorders is limited, and screening for them evidently isn’t his first priority. Because of this, the roles of both the FME and the AMHP are narrowed to “brief interviews…rather than conducting…full mental state examination[s] and the delays in acquiring the services of an AMHP act as a disincentive to initiating assessments under the MHA”. The balance between health and justice is tested, “when potentially conflicting priorities converge, and agencies struggle to put their need to deliver punishment…treatment and rehabilitation in an agreeable and appropriate order”. This is certainly something the Government will have to grapple, and address immediately if national implementation is to be successful.

There are arguments that these Sections are disproportionately used against ethnic minorities: “there ha[s] been considerable controversy…particularly concerning use of the power[s] in cases of members or minority ethnic communities”. A study by Singh et al suggested that black people are 3.83


times more likely to be detained than Caucasians, while a 2005 study by the Sainsbury Centre for Mental Health suggested that “people from ethnic minority groups are up to six times more likely to be detained under the MHA than white people”.47 While the figures are diverse, the fact remains that black and ethnic minority communities are still overly represented in the MHS and CJS.

Brought into the sphere of diversion, figures illustrate that “black communities are 40% more likely than average to be referred to mental health services through the CJS”.48 As Bourne discovered, police officers are “prone to associating black people with risk factors ...with the result that black people are more likely to be detained by police under Section 136 of the MHA and taken to a “place of safety”.49 Combined with the overt fact that the numbers of black ethnic minority communities are latently increasing within the prison population, NACRO describes the “double jeopardy” scenario. That is, diversion in this sense effectively involves a patients “transfer from one potentially discriminatory and damaging system to another” without paying due regard to their primary need for medical treatment. These figures are damming to say the least, and their discrimination does not stop there. These discrepancies come from an increased association of risk with black patients, and despite attempts by the Social Exclusion Unit, and the Care Quality Commission’s annual “Count Me In” census of ethnicity in mental health, there has been “little impact on the disproportionate number of black people...subjected to compulsory treatment”. The Reed Report’s 1994 4th Volume50 explicitly stated that for diversion to work indiscriminately for a potential benefit “it must provide a service which is appropriate and acceptable...involving staff who are sensitive to the issues of race and culture”.51 The number of recommendations made in this area, from the National Institute for Mental Health, Department of Health, the Home Office and all their caveats are plentiful. But as the latest “Count Me In” census publication

50 “Race, Gender and Equal Opportunities”.  
shows, things have not materially changed at all\textsuperscript{52} even after the unnecessary death of David Bennett in 1998, which prompted the initial cry for change in this area.

Because diversion is a multi-agency goal, the police and the health service have to work in co-operation to address this massive inequality. Funding and effective training must be invested not only to aid officers on the front line in identifying a genuine cause of concern within an individual potentially suffering from a mental health disorder, but also to provide insight into different cultures on which these communities are based in the hope that discrimination against them at least diminishes, and are not founded on the assumption and stereotype that all ethnic men pose a threat of violence.

The MHA 1983 introduced remand powers and interim hospital orders under Sections 35, 36 and 38. Section 35 will allow any Crown or Magistrates’ Court the power to remand the accused for assessment of his mental condition, while Section 36 will grant power to the Crown Court only, in remanding the accused for psychiatric treatment. Figures for 2007-8 show that Sections 35 and 36 were used 145 and 17 times respectively, and provide some comfort in showing that offenders are remanded to a therapeutic establishment where treatment and care are provided, rather than remanded in prison custody which would inevitably exacerbate their illness. But compare these figures to those under Section 47 – which allows for the transfer of detainees (the sentenced prisoner) to hospital: 900 transfers in 2009-10. The rationale for when diversion occurs and why is ambiguous. For most individuals, mental disorders are not fleeting and certainly any disorder warrants adequate care and treatment. Why the discrepancy? While I will explore further the use and remits of Section 47 transfers to prison, I will note here that the Ministry of Justice, who direct such transfers, tend only to impose them towards the end of the prisoners term. This stark contrast highlights the greatest failings of diversion, in that there simply isn’t enough psychiatric provision to deal with the numbers of offenders passing

\textsuperscript{52} \url{http://www.guardian.co.uk/society/2011/apr/06/minority-ethnic-mental-health-plan?INTCMP=SRCH}
through the criminal justice system at its initial points. Habitual criminals committing petty offences will merely be cautioned, without the provision and diversion to a mental healthcare setting, and those justly warranting treatment will be sentenced nonetheless and given second grade care until the ministry fears for the public’s safety. There is no denying that "prison in-reach services are mostly focused on medication, with little or no provision for other services such as talking therapies or counselling". As Herring eloquently expressed, “people who should be receiving treatment for their mental disorders are instead being detained in prison for offences committed because they have not received the treatment that they should have” done in the community. For diversion to succeed, the provision of mental health services in the community should first be reformed and altered, before applying such provisions in the context and for the benefit of offenders entering and residing within the CJS.

“It is possible for people who are convicted...to be diverted to the mental health system if they suffer from a mental disorder”. How does the process of diversion impinge on the courts sentencing decision? Here, Section 37 Hospital Orders and Section 37/41 Hospital Orders with Restrictions provide that the offender's mental disorder need not have any bearing on his criminal conduct, but quite simply, the courts must ask whether the defendant is, or is not, suffering from a mental disorder warranting medical treatment at the time of sentencing.

Section 37 is “the main psychiatric disposal used by the courts” and “the criteria for hospital disposal can be seen as wholly medical in nature”. 

---

Following the decision in Birch, it was held that “Once the offender is admitted to hospital pursuant to a hospital order...without restriction on discharge, his position is exactly the same as if he were a civil patient. In effect he passes out of the penal system and into the hospital regime. Neither the court nor the Secretary of State has any say in his disposal...a hospital disorder is not a punishment [and] questions of retribution or deterrence are immaterial. The sole purpose of the order is to ensure that the offender receives the medical care and attention which he needs in the hope and expectation that the result will be to avoid the commission by the offender of further criminal acts”. This statement upholds all values and objectives contained in the practice of diversion. As the treating psychiatrist maintains his position of authority, in granting leave and detention, the offender will be given care, treatment and rehabilitation for his disorder.

Section 37 can also be applied with restrictions, by virtue of Section 41. Here, “patients who have committed criminal offences can be made subject to indefinite restriction orders” by courts paying due regard to the “nature of the offence; the offender’s past history; the likelihood of re-offending, and of any risk to the public”. Thus, in comparison to a Section 37 hospital order, which is “wholly medical in nature”, a hospital order with restrictions is “a provision based primarily upon considerations of public safety”. Release is only granted and determined by the Ministry of Justice or independent mental health tribunal, and the treating psychiatrist is deprived of the power “to give patient leave in the community”.

---

The problem manifesting here are the numbers of granted hospital orders without restriction, in comparison to those with restriction. Numbers of Section 37 hospital orders have virtually halved since 1966 where 1,440 were made. In 2007-8 there were 352 hospital orders without restriction compared to 455 with restrictions. Perhaps the issue at hand lies with the power to detain and powers to grant leave of absence, transfer or discharge. The Ministry of Justice may fear the early discharge of a MDO who may reoffend, and as Section 37 hospital orders deprive them of the legitimate authority to uphold or renew detention, hospital orders with restriction evidently become the favoured method of disposal. However, the Court of Appeal’s decision in \textit{R v A}\textsuperscript{65} held “that the risk that an offender might be released prematurely by a MHT is not a ground for a court to pass a life sentence rather than making a restriction order”\textsuperscript{66}.

This preference translates and transcends a deeper meaning of the control and treatment of MDOs. Firstly, while \textit{Birch}\textsuperscript{67} upheld the paramount position of the treating psychiatrist under Section 37, the court of appeal went on to state that “the sentencing court is required to assess the seriousness of the risk not that he [the offender] will reoffend, but the risk that if he does, the public will suffer serious harm”.\textsuperscript{68} This form of thinking incubates precisely what diversion does not primarily concern or purport. The detention of an individual, through the process of diversion from the CJS, is justified through the wish to treat them, alleviate their symptoms in a healthcare setting which is undeniably better placed that its prison counterpart. Diversion does not concern detention for the possibility of risk of reoffending, and detention under the MHA cannot be abused in this way. While there will undoubtedly be circumstances where legitimate cause will justify detention because of risk: public protection cannot be the sole justification. This was demonstrated in \textit{R v Osker}\textsuperscript{69}, where the Court of Appeal

\textsuperscript{65}[2005] EWCA Crim 2077. CA.
\textsuperscript{66}Fennell, P. “Mental Health: Law and Practice”, (2011) 2\textsuperscript{nd} Edition, Jordan Publishing Ltd, Chapter 7 at page 46.
\textsuperscript{67}R v Birch 1989 11 Cr App R(S.) 202.
\textsuperscript{68}Fennell, P. “Mental Health: Law and Practice”, (2011) 2\textsuperscript{nd} Edition, Jordan Publishing Ltd, Chapter 7 at page 46.
\textsuperscript{69}R v Osker [2010] E.W.CA Crim. 955.
quashed the defendant’s hospital order with restriction without limit of time as “she posed little danger to the public, only a risk of self-harm...which did not substantiate sufficient risk to the public to justify the order”.70 This example serves to re-iterate my previous point on risk of violence and risk in general. The orders and sections serve to facilitate care provided to these MDOs, and to detain, or to continue to detain them for not, in effect, adhering or conforming to social norms is detrimental to the purpose of diversion.

The Court in Birch71 provided Guidance (at 215) for when “prison might be chosen as an alternative for a patient found to be suffering from mental disorder. The first is where the offender is dangerous and there is no suitable secure hospital bed. The second is where there was an element of culpability for the offence which merited punishment”.72 It does not aid the implementation of diversion where the number of psychiatric beds in England and Wales have dwindled since the 1960s, and where retribution warrants detention regardless of the offender’s health. This conflicts with Lord Bradley’s Report wholeheartedly, which “aims to circumvent some of the delays in dealing with mentally disordered people or to prevent them entering the criminal justice system in the first place”.73

However, there is a fundamental positive to the sentencing of an offender to a hospital order with restriction. While the treating psychiatrist loses his clinical discretion to give leave to the patient without the consent of the Ministry of Justice, “it provides for the conditional discharge of patients”.74 That is, the patient upon discharge must comply with the conditions imposed by him by the granting authority (the MoJ or the MHAT). "A conditional discharge from a restriction order in effect constitutes a strict form of compulsory treatment in

70 R v Oskar [2010 E.W.C.A Crim. 955.
the community. The necessity for such an after care provision in a given case is a common reason for a psychiatrist to recommend to a court that a restriction order be imposed. If the practice of diversion is in hope to treat MDOs, while coercive, the restriction order can in fact be used as an aid to integrate the offender back into society while still providing him with the after-care needed to ensure that his health care needs are being met and maintained. Conditions can range from maintaining contact with a social worker to taking medicine or “attendance at a psychiatric clinic”. For the Government’s national implementation to work effectively, diversion and the post-care given by the restriction order could provide the crux to keeping such offenders out of the CJS, and cared for in the MHS.

Because mentally disordered offenders can be transferred from the prison penal system at any point, Section 47 of the MHA allows for the transfer of a convicted “sentenced prisoner to hospital by the Ministry of Justice’s direction”.

There have been many significant changes in this area of diversion policy. “Since 2003 the number of in-patients transferred have begun to rise significantly, the total for 2005 being 561” and the total for 2009/10 around 900. The effect of this manifests through the increasing recognition of a genuine mental disorder warranting detention. Furthermore, following the European Court of Human Right’s decision in Jean-Luc Riviere v France, it was held the “continued detention in prison of a seriously mentally disordered person constituted a breach of Article 3”. When detaining a MDPO, it was deemed to entail “particularly acute hardship and cause[s]…distress or adversity of an intensity exceeding the unavoidable level of suffering inherent in detention”.

This, as I have mentioned previously, this does not correlate with stated practice of Section 47 transfers nearing the end of the offenders sentence. \( R(D) v \)

---

75 Application No. 33834/03.
77 Application No 33834/03.
Secretary of State for the Home Department and the National Assembly for Wales\textsuperscript{78} concerned the legality of a transfer nearing the end of the offenders sentence. While the presiding judge upheld the transfer as lawful, he did expound that there was a real sense of injustice. “Once prison serve have reasonable grounds to believe that a prisoner requires treatment in a mental hospital, the Secretary of State is under a duty expeditiously to take reasonable steps to obtain appropriate medical advice, and if advice confirms the need for transfer, to take reasonable steps within reasonable time to effect transfer”.\textsuperscript{79} The Mental Health Act Commission have long “expressed their continuing concern that transfers late in the sentence are distorting mental health law by using it for primarily public protection purposes, and that transfer should take place as soon as is therapeutically indicated rather than being delayed until risk is the primary factor”.\textsuperscript{80} By entering the MHS at such a late stage the offender, while spending years incarcerated and most probably on a cocktail of anti-psychotic and neuroleptic medication, is deprived of opportunities and schemes within the CJS of earning qualifications and learning trades which aids in his effective rehabilitation and provides him with greater life prospects once he leaves the CJS.

“Some grounds for optimism exist, although there are particular problems in implementing change at a time of financial austerity”.\textsuperscript{81} Both the NHS and the CJS face massive cuts to their funding, and it is the smaller initiatives, of mental health provisions and in-reach prison schemes which will fare worse. Continuity of care following release will be a rare commodity for the luckiest of offenders, that is, if they have been given any care at all. Figures are diverse – whether it be 70\% or 90\%\textsuperscript{82} of “sentenced prisoners [who] have identifiable mental health problems”, funding has to attack both systems detriments before embarking on

\textsuperscript{78} [2004] EWHC 2857.
\textsuperscript{79} Fennell, P. “Mental Health: Law and Practice”, (2011) 2\textsuperscript{nd} Edition, Jordan Publishing Ltd, Chapter 7.
\textsuperscript{82} Sainsbury Centre for Mental Health 2009.
such a widespread national network. The legal mechanisms in place are effective in practice, but should be used with greater care, insight and understanding of their purpose, their aims and what they can potentially deliver for the MDO. These “comprehensive plans for future service development” have incessantly been established and re-established again through government sponsored reviews and it is difficult to envisage a successful partnership between two systems reaching their maximum capacity points.

CELYN COOPER

---

Policing the Internet: The ‘Safer e-Communities’ Model

RICHARD HANSTOCK

Chapter 1 – Setting the Scene
Notes on scope, method and terminology

The purpose of this dissertation is to apply the core tenets of policing literature to the regulatory challenges of the Internet. Identifying the barriers to policing inherent in the nature of the Internet, the case for global nodal governance is made, and a template for digital policing is provided and defended.

A recent Parliamentary briefing paper (Downing 2011) heralds the development of a new National Cyber Security Programme, in consultation with NATO, the US, Europe, and other nations. This is an area likely to rise to the forefront of political debate within the next 12 months, and is thus ripe for research and commentary to inform and involve an increasingly technically literate public. Although consensus has emerged that the ever-developing digital ‘threat’ requires a comprehensive policy response, there has been little discussion of the form this ought to take. This paper aims to stimulate this debate by uniting the maturing literature on the regulability of the Internet with studies of policing as it is presently understood, pointing out the limitations of current knowledge and proposing a way forward.

Chapter 2 offers an overview of the core tenets of policing literature, followed by a tour of cyberregulatory theory in Chapter 3. Chapter 4 brings the two together, proposing a new model for police co-operation in cyberspace.

Policing studies generally focus on patrol, yet there is no demand for reassurance or order maintenance online, as communities empowered by software interacting at a distance are better equipped to self-regulate than their offline equivalents. By applying the conclusions of policing studies to the qualities of computer communications, it is argued that global interconnectedness demands
a nodal view of policing based upon a networked theory of social regulation. ‘Safer e-Communities Teams’ are proposed as a template for a new international effort to police the Internet, tasked with partnership working, public engagement and covert patrol of Cyberspace.

International co-operation and co-ordination is a recurring theme in this paper. Sovereign legislatures will inevitably disagree on matters of policy, though networked regulation will be ineffective if it is inconsistent between states. The Internet must be seen as an interjurisdictional matter, not unlike environmental and economic policies that depend upon state collaboration to succeed. Police forces, as arms of sovereign states, will be referred to as regulators, seeking to influence and improve behaviour online. The distinction between ‘policing’ and ‘the Police’ is explored in section 2; this essay adopts a broad view of ‘policing’ that encompasses much more than merely the emergency maintenance of order.

It is not the purpose of this dissertation to review tactical or technical approaches to tackling specific cybercrimes: this would confine the inquiry to institutionalised police forces, ignoring one of the central messages of this paper that our conception of ‘the Police’ is too narrow to be effective in the networked environment. The nature of cybercrime is explored in section 3, but it is meant to encompass any harmful activity mediated in some way by technology. The objective of policing on the Internet is to mitigate against these constantly developing harms; what follows is a schematic for this mitigation.

**Chapter 2 – The Thin Blue Line**

Perhaps the most helpful guides to the vast body of policing literature come from two of its foremost contributors, Newburn (2005) and Reiner (2007, 2010). These texts in particular furnish students of policing with an invaluable map of the core tenets of the subject. It is the aim of this chapter to set the scene for reflection upon these findings in the context of the information society in section
4. These findings are interconnected and largely defiant of separate analysis, but are divided into four sections for the purposes of this paper.

2.1 Policing vs. the Police

The policing function is not carried out by organised police forces alone.

There exists a curious preconception that the police are necessary to social order. Reiner (2010) terms this “police fetishism”, pointing out that institutionalised police forces are relatively new in historical terms. Response, detection, prevention and reassurance are four core police functions, but the causes of crime are best tackled in partnership with other agencies. Crime is a sociological phenomenon and the police alone cannot control it: Reiner coins Rogers’ (2002) term ‘liddism’ to describe policies that neglect the structural causes of crime. For the Metropolitan Police Service, the ‘Together’ initiative is a key part of modern professional policing, featuring in the strapline, “Working together for a safer London”. Policing is an activity undertaken by more than ‘the Police’ alone: policing partners include a whole spectrum of public and private bodies, including social services, private security, specialist investigators, and so on (Crawford 2003). Importantly, a great deal of policing is performed by ‘the internal policeman’: the super-ego, one's own conscience (Durkheim 1893).

This broad sense of policing lends itself to depiction as a social control network, the nodes of which share intelligence and resources to achieve “security governance” (Johnston and Shearing 2003). Chapter 4 will argue that this shift towards a networked understanding of policing is crucial to efforts to police the information society. Walker (2000) applauds the expansive view of policing taken by modern scholars, encompassing nodes in a “policing continuum” spanning Home Office forces and their Special Constabularies, prosecuting authorities such as HM Revenue & Customs and the Health and Safety Executive, civil servants such as probation officers, social workers and youth workers, private security professionals and their regulating bodies, commercial entities, citizen policing schemes such as the Neighbourhood Watch, and many more besides. Chapter 3 introduces some of the bodies concerned in policing the
Internet, and Chapter 4 suggests how these actors fit in to the regulatory landscape.

Walker points out that although acceptance of such a broad continuum of policing partners avoids the conceptual pitfalls of functional and instrumental essentialism, casting the network too widely may create a “serious regulatory deficit”. A comprehensive model of policing in cyberspace must operate within a framework of best practice, presenting a united front in the face of common goals, acting in a manner compatible with individual rights and the rule of law. Close involvement with the ‘extended police family’ is a cornerstone of modern police governance, concerned to work in partnership with neighbours sharing the policing landscape.

Drawing on Foucault’s 1978 lectures, Valverde (2008) comments that social power relationships, including technologies of power in policing, enable and are constitutive of practices shared by several institutions. Nowhere is this truer than in a widely-drawn policing landscape. Further, each individual-police encounter shapes the quality of the relationship between governors and governed. To the extent that these encounters can be improved through regulation, deficiencies in policing regulation have constitutional consequences, even when their targets are non-public policing bodies.

2.2 The secret social service

Society demands much more from its police service than crime control alone.

Police work can be organised into two broad categories: proactive and reactive. Common perceptions of the police tend to revolve around the latter, dominated by images of ‘blues and twos’ response policing and decisive enforcement action. Given that this is the most visible element of police work, it is hardly surprising that this has come to epitomise the popular conception of the institutionalised police force.
The role of ‘the Police’ is rather broader than is commonly perceived. Cumming et al. (1965), Wilson (1968) and others found that patrol officers are occupied significantly more with the maintenance of order than with law enforcement. This view rapidly gained broad acceptance: summarising the scholarly position in 1979, Cordner wrote, “anyone suggesting within informed circles that policing is centrally concerned with crime and law enforcement would be summarily dismissed as neanderthal.” Cordner and others, such as Shearing and Leon (1977), take issue only with the methodologies employed and generalisations made by the various studies underlying that assertion, not their conclusions.

Early studies tended to focus on patrol, supporting a conception of the police as an emergency social service, “the social equivalent of the AA or RAC patrolmen, who intervene when things go unpredictably wrong and secure a provisional solution” (Waddington 1983). A great deal of police work aims to manage conflicts and reduce risk; only in a minority of encounters do officers deem it relevant and appropriate to enforce the law and use police powers. The police are described as the “secret social service” (Punch 1979), as officers are called upon to act as “philosopher, guide and friend” (Cumming et al. 1965). A recent police foray into the microblogging arena confirms these classical findings. In October 2010, Greater Manchester Police ‘tweeted’ a live summary of each of its emergency calls over a 24 hour period (GMP 2010): a relatively crude analysis by the Manchester Evening News revealed that approximately one third of the calls taken can be described as “social work”, supporting calls from Chief Constable Fahy further to integrate ‘the police’ with other ‘policing’ professionals such as probation staff and youth workers (Scheerhout 2010).

This is certainly an accurate reflection of the character of ‘response’ police work. However, the study by Cumming et al. revealed that policing has a dual nature: control versus support. Banton (1984) develops this into a distinction between ‘peace officers’ and ‘law officers’, arguing that patrolmen are typically ‘supporters’, whilst specialist officers preoccupied with enforcement are best described as ‘controllers’. 
Shearing and Leon (1977) argue that the distinction between ‘social service’ and ‘law enforcement’ police work is unhelpful as the symbolism of the police warrant stands in the way of even amateur social work. As Reiner (2010) puts it, “[t]he craft of effective policing is to use the background possibility of legitimate coercion so skilfully that it never needs to be foregrounded”. The police role is certainly highly symbolic, representing government, law and the constitutional order; but this argument presupposes that police symbolism inspires cooperation. In fact, a uniform or warrant card can aggravate a situation as easily as it can defuse it, and part of successful police work concerns knowing when to walk away. Further, the advice and insight offered by an officer in his capacity as “philosopher, guide and friend” may be authoritative, but adherence to it is by no means compelled, especially when it is dispensed outside confrontation. Legitimate coercion is irrelevant to some encounters – consider for instance the patrol officer on the West End supplying tourists with directions and posing for a photograph. Police work is highly situational: online, even visible police work will uncover that the police brand is an impotent symbol, due to the pseudonymity of the Internet, users’ sense of detachment from the physical world in which the police monopoly on force resides, and the borderless, multijurisdictional nature of Internet communication.

In Cyberspace, ‘peace officer’ functions tend to be delegated to private superusers, able to support geographically dispersed communities by manipulating digital environments, shaping and moderating digital expression and leading and influencing opinions. Such powers are not held by states, except by court order in the relevant jurisdiction. Despite this reversal of power, all communities can and do fail to self-regulate. It will be argued in Chapter 4 that the Internet requires more than merely specialist police ‘controllers’: instead, the Internet ‘peace officer’ must focus on working internationally with partners to achieve coherent regulation across sovereignties.

Fortunately, today even specialist units are not as preoccupied with enforcement as Banton suggested. The Policing Pledge (Home Office 2008) and revised Victim’s Charter (OCJR 2005) cemented community-focused policing into the
operational values of the ‘new’ police. The Safer Neighbourhoods initiative in particular created a ‘community’ career specialism for police officers, focusing exclusively on support for local order maintenance problems and priorities. Policing with the consent and active co-operation of diverse communities pervades police organisation (Home Office 2005): specialist units such as Trident¹ are structured around active input from Independent Advisory Groups drawn from local communities, and are supported by dedicated Community Engagement Teams facilitating dialogue between specialist officers and members of the public. Other units use Community Support Officers, Special Constables, partner agencies and centrally-tasked officers to support specialist operations, providing uniformed reassurance and ‘boots on streets’ to engage with members of the public around local concerns. Winning and improving public confidence in the police and its partners plays a key role in securing public co-operation at street level, during investigations and in court, thus even ‘law officers’ have a ‘peace officer’ role. Both community-minded order maintenance and law enforcement values are inherent in policing. Thus, whether Internet policing responsibilities are delegated wholly to specialists or shared across all officers and police partners, a mix of enforcement and engagement is crucial to ensure communities are aware of their responsibilities online and receptive to efforts to enforce those responsibilities.

It is worth noting that enforcement activities offer readily quantifiable results, with detections, arrests, tickets and so on all contributing to the ‘thin blue line’ delusion. Proactive ‘social work’ and engagement activities are less quantifiable however, and are thus at risk of being compromised or discontinued in the wake of police spending cuts. A crucial task for police leaders therefore is to appreciate the special importance of public engagement in the reduction of cybercrime, recognising and rewarding achievement and innovation in projects intended to improve the regulability of cyberspace.

¹ Trident is the specialist gun crime unit within the Metropolitan Police Service, tasked with the prevention and investigation of shootings across London and gun-related murders within London’s black communities (Metropolitan Police Service 2011).
Sykes (1986) offers a moral defence of ‘peace officer’ patrolmen, suggesting that all police actions – including those which indicate dominion (Marenin 1982) – are expansive of, not threatening to, human freedom. According to Sykes, disorder undermines community life, thus “[i]n responding to the mandate for order maintenance, the police create a sense of community that makes social life possible.” The following section on police discretion presents an ethic of responsibility as key to good policing, but according to Sykes the police must also embrace a “moral leadership role”. The extent to which the police are crucial to social life is deeply questionable: per the above, the police are not principally moral arbiters, but emergency maintainers of order; thus, it is submitted, the police facilitate social life.

All police social work is amateur: the division of labour in rehabilitative criminal justice provides social workers and probation officers to those identified as being in need of moral leadership, working with parents, teachers and friends to develop key socio-moral values. Morality in urban, non-urban and digital spaces is diverse, personal, familial and decentralised, and part of the challenge of modern policing is to understand, respect and embrace this diversity, recognising the limits of the police role and the value of sharing information with partner agencies.

Later in this dissertation it will be shown that programmatically empowered digital communities tend to self-regulate without police input, the order maintenance function being performed by superusers and volunteer moderators performing an overt, rather than a secret, social service. Nonetheless, it will be argued in Chapter 4 that the Internet peace officer should play an important role in the oversight of private order maintenance, working with norm-enforcing private actors to identify criminality and broaden respect for law in cyberspace.

---

2 Shearing and Leon (1977) argue that the police warrant militates against the ‘social worker’ conception, owing to the omnipresent “background possibility of legitimate coercion” (Reiner 2010). The counter-argument that social workers may also wield similar powers leads us down a cul-de-sac: the point is that ‘social work’ is intrinsically part of the police mandate.
2.3 The good policeman and the rule of law

The police routinely under-enforce the law.

In *Police: Streetcorner Politicians*, Muir (1977) uses Weber’s (1919) model of the ideal politician to come to a view of the qualities of a “good” police officer. A “mature man”, the good police officer is endowed with both passion for his cause and perspective in his means having embraced the “ethic of responsibility”.

“Intellectually, [the good policeman] has to grasp the nature of human suffering. Morally, he has to resolve the contradiction of achieving just ends with coercive means.” Justness on this account is highly subjective, and much of the policeman’s lot is the pursuit of justness “in spite of all” competing pressures: in exercising their relative autonomy (Cordner 1979), police officers choose which values underlie their notions of justness. This is an unavoidably individualistic process, begging the questions “whether we can rely on the personal sense of justice of patrol officers, and how its exercise can be made accountable” (Reiner 2010). Broad discretion is inherent in the police licence, but the values at stake are so crucial to what Fuller called in 1964 “the inner morality of law” that a framework of consensus and safeguards has evolved to structure and restrain that discretion.

Muir’s account helps us to understand the street-level reconciliation of coercive means with just outcomes, and offers considerable insight into the scope of police action and the dangers of miscalculating justness. Police work is highly situational, thus “discretion… cannot be avoided” (Marenin 1982). Unsurprisingly, the structuring of police discretion is one of the key concerns of police governance, perhaps the best example being the introduction of positive arrest policies in cases of domestic violence following the Minneapolis Experiment (Sherman and Berk 1984) and in response to pressure from feminist groups. Attitudes to discretion vary between jurisdictions, over time, in urban and non-urban contexts, during times of social unrest and public emergency, and so on. Generally, civil jurisdictions tend to be less permissive of discretion in the administration of justice than are common law traditions; the policy of zero-tolerance policing will be discussed in the following section. Given these
differences in governance, it will be argued in Chapter 4 that interjurisdictional police co-operation requires an intrinsically permissive, if not a consistent, attitude to discretion in the allocation and application of police resources.

As the late Lord Bingham (2010) put it, “[q]uestions of legal right and liability should ordinarily be resolved by application of the law and not the exercise of discretion”, and “[n]o discretion should be legally unfettered.” Discretion, tempered by proportionality and accountability and supervised by the courts, must pervade all areas of policing, to avoid what Skolnik (1966) called ‘justice without trial’. Officers are trained to be aware of their prejudices, address them, and ensure they do not interfere with their duties – though even recent observational studies (e.g. Loftus 2009) suggest the reality is somewhat departed from this ideal. Lawyers, rights and courtrooms comprise the first line of defence against arbitrary discretion, and such damning reports as Macpherson’s (1999) highlight the many dangers of the inequitable application of coercive power.

Race is an important part of the story of modern policing and the quality of the relationship between citizen and state, but given the (ever-diminishing) degree of economic power necessary for access to computer networks, questions of social class are more relevant to the focus of this dissertation. Loftus (2009) points out that class oppression has been somewhat replaced in political discourse by matters of race and individual equality more broadly. Her observational study revealed that police suspicion was directed mostly at poor white men, calling for a scholarly refocus on matters of class, across all ethnic groups.

In response to presentations of the police as an instrument of class repression, Marenin (1982) points out that owing to the breadth of their mandate the police in aggregate produce both specific and general order, the interests pursued being entirely dependent upon the circumstances of individual encounters. Marenin notes that policing “is done largely by members of the working classes” to support his argument that the police cannot be used to impose specific order
in contradiction of their own interests. It will be argued later in this dissertation that digital communities tend to achieve general order by self-regulation more effectively than their offline counterparts, and that enforcement activity online will generally offer less scope for discrimination of all varieties, curbing opportunities for specific dominion quite separately from the class makeup of police forces. Chapter 3 will point out the ‘digital divide’ between those with and without access to the Internet. Though policing is said to target the most vulnerable, the most vulnerable lack access to the Internet, albeit decreasingly. The specific order impact of policing the Internet is therefore likely to be similar to the specific order impact of traffic policing, which by its nature focuses on the motoring public, who necessarily enjoy some degree of economic power: this is not discrimination, at least not in a normative sense, determined instead by the ‘resident’ or ‘available’ population (Bowling and Phillips 2007) of online communities, a structural matter of access and lifestyle rather than a product of deficiencies in police decision-making. These arguments will be developed further in Chapters 4 and 5.

Another of Loftus’ findings was that the police remain resistant to change. This frontline inertia occurs as educated ‘management cops’ generate and embrace change which does not filter down to street-level practice (see Holdaway 1977 on managerial versus practical professionalism), for reasons deeply rooted in ‘cop culture’ (see Skolnik 1966 and Reiner 2010). Organisational culture is the “lens through which... the shape and contours of the external world are fixed and established as reality” (Manning 1982), thus it is frontline identity that is the source of this inertia. To the extent that frontline policing is a blue-collar occupation concerned mostly with the lower social classes, management officers are white-collar workers with a broader, now classless, view of the police role.

3 There is potential for class- and even race-based suspicion in traffic policing, with suspicion directed at older, poorly-maintained vehicles: see Loftus (2009) on the highly aesthetic nature of police suspicion. On the Internet, it is possible to determine the basic configuration of a user’s system, for example the age of the operating system or the revision of the Web browser being used, as well as its approximate geographic location. Traffic police can at least root suspicion of older vehicles in an objective concern for the detection of vehicle defects and the public interest in roadworthiness on the public roads network, yet as there is no legal conception of ‘Internet-worthiness’ it is not unlawful ignorantly to connect a virus-ridden computer to a communications network. If such a concept developed, policing online may come disproportionately to affect those who appear less technologically literate.
Officers specialising in Internet policing are likely to be more educated and thus less likely to be members of the working class, more receptive to change and less likely to act against the interests of the more affluent digital classes. Loftus calls for improved, theoretically sophisticated training to eradicate class stereotypes from the police mindset, and Chapter 4 argues that this training should incorporate the many issues relevant to efforts to police the information society, including an awareness of the state and trajectory of the digital divide. However, storytelling is central to the fabric of police culture, perhaps more so than classroom training, as there is a view on the front lines that ‘everything is perfect in training land’. Seasoned officers with professional, confidence-inspiring stories should be empowered to act as role models for newer generations of police recruits.

Policing online is a new effort however, so it is training that will lay the foundation for the stories that are to follow in Internet policing. In his review of police leadership and training, former Chief Constable Peter Neyroud (2010) proposes the creation of a chartered professional body for policing, charged with, *inter alia*, the development of a rigorous national ‘fitness to practice’ qualification supported by structured continuing professional development. This change presents a prime opportunity to inject and reinforce both competence in digital policing and rule of law values on a national level, and could address frontline inertia across all areas of policing. National accreditation ensures that digital expertise is accessible to all force areas to a consistent standard, avoiding postcode lotteries in policing quality. Neyroud envisions that this body will use its influence to draw international membership, and could become a steering group for precisely the kind of international police co-operation it is the object of this dissertation to promote.

This section has shown that although police discretion is inevitable, its exercise is bound up with the ethic of police responsibility which defines the good police officer. Part of this ethic is to accept the division of labour in criminal justice between police officers, advocates, judges, and specialists such as probation officers and youth workers. Police officers must ensure deference to law and
adherence to the rules of evidence, acknowledging the difference between suspicion and proof and embracing the moral proprieties of trial, conviction and sentence. Officers must avoid the temptation to abuse discretion and popular ignorance in acting as street corner judges and juries: policing on this conception becomes a term of art, the inner morality of policing being an application of the inner morality of law. Through enhanced training, supervision and competent governance, police forces globally can adapt to the challenges of the digital environment whilst combating arbitrariness and improving effectiveness in all quarters, enforcing the law while upholding the rule of law.

2.4 Policing and crime control

The police are marginal to social order.

For Sir Robert Peel, ‘the test of police efficiency is the absence of crime and disorder’. Policing by consent demands ‘public approval of police actions’, and the public favour ‘more police on the streets’. Yet, patrol has little effect on crime rates, tending to displace and discover, rather than prevent and deter, disorder. Kelling et al. (1974) conducted the famous Kansas City patrol experiment, which found no relationship between foot or mechanised patrol and policy objectives including crime rates. Although this study was subject to strong methodological criticism, the results are unsurprising: the chances of even a plain-clothes officer patrolling randomly being in the right place at the right time to catch and apprehend a criminal ‘in the act’ are very small indeed. Intelligence-led deployments may be more successful, but do not solve the problem of temporary displacement of crime and criminals.

Preventive patrols always carry a reassurance agenda, contrary to Peel’s view that that visibility is not determinative of police effectiveness. Public confidence is now the principal measure of police performance, and confidence fetishism fuels reassurance patrolling: Commissioner Sir Paul Stephenson stresses the paramount importance of “uniformed governance” of the streets (Edwards 2009), and has recruited the country’s largest Special Constabulary to expand visible police presence and thus reduce the fear of crime. Police Community
Support Officers acting under the Police Reform Act 2002\(^4\) denote a similar ‘boots on streets’ rationale. The Philadelphia Foot Patrol Experiment (Ratcliffe \textit{et al.}, forthcoming) suggests that intelligence-led fixed beats do have some effect in violent crime hotspots, even adjusting for displacement. However, it seems that routine patrol and neighbourhood beat policies which are the backbone of police presence\(^5\) are concerned with something other than crime control: the reinforcement of police legitimacy, premised, ironically, on a misconceived crime control agenda. The problems of “uniformed governance” of the international communications network, and the consequences for public confidence in the digital order, will be discussed in Chapter 4.

Jackson \textit{et al.} (2009) conclude that it is the low-level ‘social work’ function of the police which drives public confidence today, with the police held responsible for “neighbourhood cohesion” and the maintenance of strong communities. Jackson \textit{et al.} draw upon the concept of “signal crimes” (Innes and Fielding 2002) to suggest that perceptions of security and neighbourhood cohesion are disproportionately influenced by sub-criminal “signal disorders”, which are seen to be the proper focus of the police, rather than overall crime rates, fear of crime or perceptions of risk. This view supports the argument that the expansion of criminality under New Labour has created ‘liability for failure to reassure’ as a tool of the “moral police”, accountable for the protection of vulnerable autonomy and enforcement of an emerging ‘right to security’ in the maintenance of social order (Ramsay 2008). The extent to which a state perception of vulnerable cyberautonomy has lead to the overcriminalisation in the face of ‘new’ networked harms is an area of particular relevance to any analysis of cybercriminality.

\(^4\) Police Reform Act 2002 c 30.
\(^5\) See, for example, the ‘single patrol’ policy rolled out by Metropolitan Police Service, requiring patrolling officers to work alone, not in pairs or groups, in order to spread routine patrols more broadly across London (Metropolitan Police Authority 2010). Officers reacted negatively to this policy, concerned that ‘management cops’ were prioritising public confidence and reassurance above officer safety, forgetting the risks inherent in routine frontline policing and the importance of readily available backup.
Policing on this account is symbolic, not instrumental (Reiner 2007), “communicat[ing] a sense of ‘guardianship’ to members of a community” (Innes and Fielding 2002). This is fortunate, as studies demonstrate that policing has little effect on the crime rates that are the focus of police rhetoric. However, the policing of signal disorder and crime rates may be linked by the ‘broken windows theory’ (Wilson and Kelling 1982): early police intervention in minor disorder can prevent escalation towards more serious crime later on. This is the rationale behind dispersal orders, used to great effect as part of policing initiatives such as Operation Safe Nights (Heart of London 2010). Reassurance patrols combined with robust early interventions which bring about new diversionary disposals devised as part of the ‘rehabilitation revolution’ could serve both to secure vulnerable autonomy and cause a reduction in hotspot crime, provided they address the underlying causes of crime (Burney 2009). However, this is not to say that officers should abandon current faith in discretion: although ‘zero tolerance’ policing was credited with causing a fall in New York crime rates, closer analysis suggests the crime drop was merely coincidental (Reiner 2010), and overzealous enforcement of minor offences is likely to lead to a drop in public confidence, particularly among affected groups (David 2010).

Equally, mere presence is insufficient. Wilson and Kelling speculate that volunteer Guardian Angels may have a deterrent, reassuring effect in communities, but the mixed reception in the UK to employed City Guardians and even Police Community Support Officers suggests their impact is limited, with the public preferring to see ‘proper police officers’. Police officers symbolise security in the popular eye: the public are aware of the differences in legal powers and levels of training, counting even volunteer Special Constables as ‘proper police’. Police support and partner patrols nonetheless perform important community functions. They do have some reassuring effect, and may even enhance the status and presence of attested officers. Clearly, further research into public reactions to auxiliary police patrols would be desirable.
Chapter 3 explains that ‘software supremacy’, the decentralised architecture of the Internet and the pseudonymous nature of online communication mean that pseudo-public spaces online are essentially under private control; thus low-level disorder is effectively subject to private jurisdiction. Chapter 4 will argue that the ‘Internet beat officer’ ought to act as a conduit between public values and private actions, exploiting the police ‘brand’ to influence behaviour by publicly approving and accrediting compatible examples of ‘private’ justice. Echoing partnerships with offline private security groups such as the Security Industry Authority, the police should continue working with partners to inspire and approve responsible online gatekeeping. More serious crime becomes a matter for ordinary detection, coercion and prosecution, but ‘patrol’ online has a different character, and there is little demand for public police reassurance. Order underlies cohesive communities, and disorder undermines trust: though commercial solutions such as encryption and Verified by Visa provide a technical ‘target hardening’ remedy, crime is a social, not a technical, problem, requiring partners to identify and provide moral leadership to those acting in accordance with anti-social norms.

2.5 Concluding thoughts
Many popular myths exist about policing, and a considerable body of research has played a significant role in dispelling these. For instance, as Marenin (1982) explains, “[s]tudies of the police at work agree on two issues: one, that police work is discretionary; and, two, that most of what the police do is not specifically concerned with crime.” Reiner (2007) offers a rather more comprehensive account, spanning seven key findings: the police are marginal to social order; the police role is not primarily law enforcement; the police exercise considerable discretion; police work is shaped by situational, rather than legal, factors; policing mainly targets the powerless; traditional policing has little crime control effectiveness; and policing is more symbolic than it is instrumental. These findings have been distilled into four sections above, providing a workable structure for later evaluation.
Chapter 3 – The Network of Networks

This chapter provides an overview of the nature of the Internet, explaining the challenges it presents to conceptions of policing in ‘real-space’. In defining what is meant by the Internet, the physical and logical layers of the network will be explored, and the dependency between the user experience of the network and the characteristics of the underlying technology will be established. A brief history of cyberregulatory theory informs a modern understanding of the demands of policing in this unique environment, which will be defended as the framework for the conclusions that follow.

3.1 What is the Internet?
Understanding a technologically mediated social space.

As the Internet has matured, so too have perceptions and understandings of it. Early literature has an air of “gee-whiz futurism” (Ross 1999), impressing upon readers that the computer movement had opened up a brave new world intelligible only by computer experts. Certainly, what was felt were the shockwaves of a technological revolution, with its own fascinating history. This revolution placed affordable computing power and powerful telecommunications equipment in homes and offices the world over, pervading and transforming modern life to the extent that, at least for certain sections of society, the relationship between person and machine was not merely instrumental but integral to self-perception (Turkle 1995).

The crucial phase of this revolution was the linking together of these electronic powerhouses into data-sharing networks. Murray (2010a) offers a comprehensive history of the modern Internet, but for present purposes it will suffice to note that the present-day Internet, or International network, consists of an indeterminate number of local and wide area networks of computers of all shapes, sizes and applications, communicating at ever-increasing speed, serving billions of users worldwide. As a consequence of our increasing reliance on
technology, the data housed within and the physical processes controlled by these networked computers become ever more valuable.

*Cavanagh’s modes of analysis*

Given the breadth of its application, ‘the Internet’ is an analytically ephemeral term. Technically, the Internet is the medium upon which Cyberspace, including the World Wide Web, rests; though ‘the Internet’ is popularly used to conflate both infrastructure and application. Cavanagh (2007) argues that interdisciplinary research into the field has muddied the waters further, suggesting a framework for analysis which respects the social and technical structure of the network of networks as well as its broader impact and popular understanding, presenting ‘the Internet’ as simultaneously a social space, a communications medium and a technology.

(i) **The Internet as a social space**

The first of Cavanagh’s modes is rooted in the functionality of the Internet as a forum for all kinds of exchange. According to the father of the World Wide Web Sir Tim Berners-Lee (2000), “[t]he Web is more a social creation than a technical one”. The focus here is on individuals forming collectives, making friends, sharing ideas and participating in the marketplace. Cyberspace is on this conception an inviting new agora, with international membership and a broad sense of community. Users align around shared interests and engage each other in dialogue: “[t]hese days, insecure in our relationships and anxious about intimacy, we look to technology for ways to be in relationships and protect ourselves from them at the same time” (Turkle 2011). It is argued later in this dissertation that strong communities and the norms therein are crucial to effective Internet policing.

(ii) **The Internet as a medium**

The second of Cavanagh’s modes concerns the Internet as a dimension of the mass media, one particular social application of the Internet. The maxim ‘he who controls information, controls the world’ is central here, as gatekeepers and opinion leaders online inspire and influence action and discourse in Cyberspace.
A discussion of the role of Internet gatekeepers in shaping “lex informatica” follows later in this Chapter, which will support the finding that it is crucial for police forces to work in partnership with these powerful organisations if they are to be effective in influencing behaviour.

(iii) The Internet as technology

The third of Cavanagh’s modes of analysis rests upon the technical conception of the Internet as a ‘network of networks’. The particular characteristics of this technology in its present form – notably the decentralised, neutral, end-to-end, borderless topology of the Internet (see e.g. Murray 2010a) – are crucial to understanding why it is so difficult for states to influence behaviour online. The regulability or otherwise of the Internet rests upon our understanding of the technology and its capabilities, particularly whether or not it is possible to control the flow of information across borders. As Negroponte (1995) put it, “the change from atoms to bits is irrevocable and unstoppable”, with profound consequences: the network of networks opens new avenues of criminal opportunity, challenging traditional understandings of harm (McGuire 2007).

The rise of self-publication and citizen journalism, made possible by the global span of the decentralised network, enhance discourse, stimulate innovation, enrich digital footprints and threaten user privacy. The infinite replicability of bits challenges our notions of property, theft, and even trespass, while technological convergence invites access to a wealth of information in the palm of our hands (Murray 2010a). These social consequences are inseparable from their technological roots, a proposition central to the arguments in this dissertation.

The Internet is, of course, simultaneously each of these three: a social space, a mass communications medium, and a network of networks. Nonetheless this analytic division is a useful illustration of the characteristics of the Internet, an appreciation of which serves to illuminate the model that I propose in Chapter 4.
“Homo digitalis”
Without users, the Internet would be just a technology, not a social space: this space cannot fully be understood without a picture of those who use it. For Katz, writing in 1997, “[t]he Digital Nation constitutes a new social class”, unrepresentative of the wider population, disproportionately “young [...] affluent[,] richer, better educated, [...] white[,] libertarian, materialistic, tolerant, rational, [and] technologically adept”. Since then, “homo digitalis” (in dog Latin) has enjoyed rapid diversification and enjoys a user base much more representative of the population as a whole. In the 2009 Digital Anthropology Report, Zeitlyn stratifies the cybercitizenry into six distinct “Tribes”, according to technical literacy and their use of the Internet, spanning “Digital Extroverts”, “Social Secretaries” and “Timid Technophobes”. If this is the future of class in the information age, groups will be distinguished not by their physical characteristics but by the choices they make – and their capacity to choose.

Perhaps a more telling approach is to track the ever-closing “digital divide”, between those with Internet access and those without.6 Although participation in the information society continues to widen and deepen, the older and less educated are more likely to lack Internet access, even within Europe. Cavanagh (2007) argues that whilst ever this is true, regulators must satisfy an intelligentsia able to critique the social order – a far cry from the paradigm poor and ill-educated ‘usual suspects’ that dominate our prisons and are so often the objects of police suspicion offline.

Hollinger (1997) offers an impressive collection of essays that tell a story of the discovery and criminalisation of cyberharms and the social demonisation of hackers. Public and judicial perception of the deviant intelligentsia is largely negative: following Evan Davis’ bemused interview of 22 year-old hacktivist “Cold Blood” on Radio 4’s Today programme, a pejorative e-mail from a listener was read out speculating whether the interviewee “had a girlfriend” (Bowbrick 2010). Computer hackers are popularly perceived as socially awkward

---

6 Projects examining the breadth of the digital divide tend to be concerned with e-democracy and social inclusion generally (e.g. Hill et al. 2004), but are nonetheless useful here.
‘computer nerds’, yet this is likely to change as the general level of technical literacy improves. The motivations of those drawn to cybercriminality are equally as broad as those who commit crime offline: whether committed for self-satisfaction, technical challenge, a need for peer respect, sexual gratification, criminal gain, commercial advantage, private vengeance or for political reasons, the perception of anonymity online and the tradition of pseudonymity amongst Internet users is unlikely significantly to change (Wall 2007).

International public policies which further “real access, reach and socially responsible connectivity” (Hill et al. 2004), together with education in the meaningful, creative and socially constructive use of these technologies, is paramount, and could be coupled with ‘netizenship’ programmes. Thus, regulatory attention to law, education and markets – including, it is submitted, by the police – could help to shape the norms of behaviour online, strengthening digital communities which, as will be shown, tend naturally to self-regulate.

3.2 Regulating behaviour on the Internet

*Networked governance: central governments and the decentralised communications environment.*

Following on from the previous chapter, policing can be defined as the promotion of pro-social behaviour through a range of formal and informal, physical and non-physical controls. As digital interaction has matured, so too have our understandings of the nature of Cyberspace: this chapter charts the ascent of Internet scholarship towards the model upon which this dissertation is founded.

*The new home of Mind*

The conception of the Internet as a social space underpins the early vision of Cyberspace as a *place*, an extraterritory, above and between sovereign states. One of the most interesting studies into the social impact of the Internet concludes that the relationship between man and machine has become so blurred as to have inspire new postmodern understandings of the self (Turkle
1995). It is this intense immersion of technology into personhood that underlies the perceived separability of Cyberspace from the physical world, which led to John Perry Barlow’s infamous Declaration of the Independence of Cyberspace in 1996:

“Governments of the Industrial World... I come from Cyberspace, the new home of Mind... You are not welcome among us. You have no sovereignty where we gather... You have no moral right to rule us nor do you possess any methods of enforcement we have true reason to fear... Cyberspace does not lie within your borders... We are forming our own Social Contract.”

The message is that the decentralised nature of the Internet gives rise to regulatory arbitrage: that there can be no law, and thus no law enforcement, in Cyberspace. Further, Barlow (1992) argues that the information society disintermediates content from carrier: the infinite replicability of bits and Cate’s (1997) four principles of data growth mean that we have lost control over the expression of ideas and have thus moved beyond the scope of traditional copyright law. Accordingly, we must learn to extract value from ideas themselves, selling “wine without bottles”.

As the father of the cyberlibertarian movement, Barlow’s sentiments are bold and widely cited, though “increasingly mocked” (Morrison 2009). The most straightforward criticism of the idea that laws cannot be enforced online is that Internet users and their hardware remain at all times physically within the coercive scope of a nation-state (Reed 2004). This argument ran the cyberlibertarian school aground, becoming the central tenet of the cyberpaternalist movement.

As the information age has matured, its focus has shifted from production to control (Murray 2010a), resulting in (primarily commercial) innovation that aims to plug the void left by this disintermediation. Digital Rights Management technologies use cryptography to restrict the use and free distribution of content
such as music and video, aiming to put a stop to unlicensed use of protected material. Innovation such as this inspired Professor Lawrence Lessig’s applauded 1995 work, *Code and Other Laws of Cyberspace*.

**Lex informatica**

In *Code*, Lessig (1999) proposed a highly influential regulatory model of the Internet centred upon the user, the “pathetic dot”. Lessig identified four modalities of regulation exerting pressure on practical reason: law, norms, markets, and architecture. “Law tells us not to do something, and if we do it, we get punished. Norms will get us shunned socially instead. Market forces will cost us money, and architecture will simply not permit us to do it” (Nehrlich 2000). Lessig argues that on the Internet, unlike in the physical world, code is the most powerful regulator. Offline, we are enabled and restrained by the architecture of the physical environment: walls, doors, desks, language and so on. The architecture of the digital space is ‘code’, the protocols and structures that shape interaction and information flows online. Joel Reidenberg (1998) coined this *lex informatica*, comparing rules for data flows and information policies to the development of *lex mercatoria* in the Middle Ages.

The argument that “code is law” is based on the technical features of the network of networks, beginning with the physical and logical infrastructure upon which the transportation, presentation and manipulation of content relies. Computer network engineers describe the functioning of computer networks by reference to a seven-layer Open Systems Interconnection (OSI) model, upon which Murray’s (2007) concept ‘layered vertical regulation’ is based. Essentially, because the application set\(^7\) is dependent upon the transport set,\(^8\) manipulation of the physical and logical infrastructure of the network shapes the content of Cyberspace by altering the capabilities of the layers above. In terms of the earlier distinction between the Web and the Internet, the former is a product of the application set, while the latter comprises the transport set.

\(^7\) The application set comprises the Session, Presentation and Application layers.

\(^8\) The transport set comprises the Physical, Data, Network and Transport layers.
Since changes in layers below affect the capabilities of the layers above, it follows that the strongest regulation occurs at the lowest levels: akin to ‘designing out’ crime in urban planning, not unlike the ‘bum-proof benches’ and ‘tramp sprinklers’ in the *City of Quartz* (Davis 1990). However, one of the founding norms of the Internet was that of ‘net neutrality’: a neutral network transports content freely and without discrimination, regardless of the nature of the traffic flowing across it. Berners-Lee is one of the leading advocates of network neutrality, concerned principally with commercial exploitation of the potential to compromise the openness of the network, as for-profit service providers compete to provide the crucial ‘last mile’ between the end-user and the broader network. Discussion of net neutrality and the responsibilities of providers to preserve it appear later in this dissertation.

Assuming that manipulation of the transport set is off-limits, the potential to shape online conduct through code in the application set remains enormous. Code sets the terms on which Cyberspace is offered. The Internet is a network of networks, each located within different sovereignties: as Cyberspace flows across state borders, code and content hosted in one country is accessible outside the jurisdiction of the host state. Further, as the Internet lacks a central authority – its architecture is end-to-end – no one body has the monopoly on access control, meaning that any two networked computers can communicate without centralised filtering or approval. This decentralised approach is crucial to the wartime resilience of the network, creating by design a network of independent nodes with no central Achilles’ heel.9

The neutral Internet is an end-to-end network spanning continents without extensive filtering at state borders. Thus, Lessig argues that relative to “real-space”, architecture is stronger and law is weaker online. Norms are weakened by lesser social accountability online, as interactions take place behind pseudonyms and a perception of anonymity; and since markets are dependent upon law and norms, their influence is weaker online too. Digital architecture is therefore the strongest modality: “code is law”.

---

9 For more on network topology and the consequences for regulability, see Murray (2010a).
Wall (2007) illustrates software supremacy with a case study of the regulation of ‘spam’, unsolicited nuisance e-mail messages. Spamming is a surprisingly lucrative activity with a high return on investment, particularly as ever-richer digital footprints enable more closely targeted unsolicited advertising. Sovereign nations reached different conclusions on the issue, enacting opposing laws: the US CAN-SPAM Act adopted an ‘opt-out’ framework on spam, while the European Union chose a diametrically opposed ‘opt-in’ approach. Despite fears that jurisdictional arbitrageurs would relocate their spamming operations to nations with no such controls, no effect on spam volumes could be determined (Kigerl 2009). Fortunately, user groups such as Spamhaus united around common norms, and continue to share information and expertise in support of the most effective remedy: spam filtering.

Spam filtering is an example of norm-based code achieving a regulatory outcome that lawmaking states failed to achieve, forming part of the *lex informatica* and playing a significant role in enhancing the quality of the online experience. However, code, like law, can be applied for less benign purposes. Echoing Berners-Lee’s concerns about commerce and neutrality, it is commercial governance of code about which Lessig was most concerned. The Internet is essentially a network of privately controlled domains: as certain domains are more popular than others, and webmasters can control the code and thus the terms on which their services are accessed, these gatekeepers must act responsibly (Laidlaw 2010). Lawmakers, robbed of digital potency, are accountable to their constituents, while empowered digital architects are accountable only to the marketplace. It is submitted that market forces alone offer too crude a safeguard for values such as network neutrality. Rather, the Internet police ought to monitor and steer private gatekeepers towards an ethic of responsibility; an argument that will be developed in Chapters 4 and 5.

---

10 Cybercrime will be discussed in the following section.
Symbiotic regulation and the active matrix

Murray (2009) develops Lessig’s theory, presenting a more nuanced thesis of the Internet as a social space, based upon “network communitarianism”. Murray argues that Lessig’s ‘pathetic dot’ is ‘not so pathetic’: users make decisions and communicate opinions, forming in aggregate an “active matrix” of communicating dots. The ‘pathetic’ dot is in fact empowered by its membership of a community, which collectively develops and enforces social norms, drives market forces, elects lawmakers, and writes its own code. The decentralised architecture of the Internet facilitates this community, enabling instantaneous communication, low-cost self-publication, citizen journalism, reader-author interaction, social and professional networking and online petitioning, on a global scale. Attempts to regulate inevitably engage the active matrix: opinions on attempts to regulate behaviour are formed, shared, developed, and ultimately communicated back to the regulator, directly, through market choices or even subversively. The active matrix chooses which technologies to accept and which to reject: therefore, the real power lies in the dialogue between regulators and regulatees.

Recognising the empowerment of the pathetic dot is to note that users are the key to effective regulation on the Internet. External attempts to regulate must therefore correspond with the norms and values of the target communities if they are to be accepted and internalised by the active matrix. Regulators must therefore map out those values by engaging with “homo digitalis” (Zeitlyn 2009), becoming part of the active matrix itself. Police forces, it is argued, in patrolling the Internet can aim to map the matrix and communicate this back to regulators.

This is not to argue however that some unified cyber-*demos* contains all optimal regulatory settlements: perfect regulation is impossible, but imperfect regulation informed by and adaptive to the norms of the target communities can nonetheless be effective. The Internet is primarily a collection of microcommunities, interest groups with their own norms and priorities. Murray (2010a) suggests that the virtual public sphere is more perfect than its offline counterpart, as Internet pseudonymity frees contributors of many factors that
chill discourse in the world of atoms. Habermas suggests the opposite, echoing the concerns of Sunstein (2001) that the Web hosts only a cacophony of speech, being too chaotic for any public spheres to form at all (Jeffries 2010). Nonetheless, it is hard to deny that participants in the information age are generally better informed and more able to participate in discourse than were their offline forbears: within the cacophony can be found some quality “constitutional dialogue” (Bakker 2008) amongst the “intelligentsia” (Cavanagh 2007) of “homo digitalis” (Zeitlyn 2009). For regulators, this dialogue offers an interface with the norms and concerns of the active matrix, and is thus the source of effective regulation online.

Recent events illustrate the active matrix in action. Upon the police closure of the Fitwatch website for offering ‘advice’ to demonstrators on Millbank during the 2010 student protests in London, the offending content had been republished on “over 100” websites by unaffiliated users (Fitwatch 2010). The active matrix, empowered by readily available self-publication tools, thus rendered the police intervention ineffective; in fact, when Fitwatch resumed service it thanked the Metropolitan Police “for giving us such huge publicity”. However, revolt against the Digital Economy Act 2010\textsuperscript{11} was less potent: despite widespread calls for repeal pending fuller debate, together with independent condemnation from the LSE of the measures intended to combat illegal filesharing (Cammaerts and Meng 2011), judicial review of the Act failed and the law remains in force. Whether this is a point about the potency of the active matrix or flaws within the legislative process, truly networked governance ought not to be so stubborn.

\textit{Channels and gatekeepers}

In conceiving of the Internet as an element of the mass media, we see the application to the online world of Lewin’s (1947) “channels and gatekeepers” thesis. The aim of network gatekeeping theory is to identify information flows in cyberspace, and those who can shape and influence them (Barzilai-Nahon 2005, 2006). Laidlaw (2010) refines this model in light of the Lessigian supremacy of ‘code’ (explored above), arguing that some Internet gatekeepers enjoy so much

\footnote{Digital Economy Act 2010 c 24.}
control over the flow of information in Cyberspace that human rights responsibilities and vicarious liabilities ought to be assigned to them, to a degree relative to the impact the gatekeeper has upon the democratic discourse of the gated. She calls this subclass ‘Internet Information Gatekeepers’, accountability of which is key to both free discourse and the equitable credentials of lex informatica. Both are crucial to the regulation of behaviour online.

In addition to Murray’s model of symbiotic regulation, Bernal (2010) identifies another symbiosis in cyberspace: between users and gatekeepers. Users provide personal information to private companies in exchange for free-of-charge access to services such as Google Mail and Facebook, which generate revenue through targeted advertising. Bernal is concerned that users undervalue the currency of their personal data in this “Web 2.5” transaction, calling for regulators to continue to protect users against the potentially significant consequences of lost or misapplied data. Bernal’s “Symbiotic Web” has important side-effects: the more personal information users are prepared to divulge, the more complete our ‘digital footprints’ become, offering not only more accurate audience profiles for advertisers, but also greater opportunities for identity theft, together with a rich vein of investigative leads for state police. Concerns about data privacy are vital if users are to continue to entrust gatekeepers with the personal information that underpins “Web 2.5” business models. Where this information is collected responsibly, kept safely and used in accordance with transparent privacy controls defaulted in favour of the user, the symbiosis between user and provider is fair. Regulators must work with gatekeepers to ensure this fairness is built in to their systems.

Markets alone are insufficient in ensuring that gatekeepers embrace the ethic of responsibility. As Bernal (2011) puts it, “dogs will be dogs”: those with the power to shape lex informatica must be “trained” to do so responsibly, by expressing legal and market decisions consistently with international norms prioritising rights-consciousness. Regulatory intervention increases the economic cost of responsibility; but those gatekeepers who wield the most power, and thus must take the most responsibility (Laidlaw 2010), are inevitably
also the most profitable, and are thus by happy coincidence the most able to afford the economic cost of socially responsible governance.

*The brittle network*

The Internet is the infrastructure of Cyberspace, the human experience within the code that flows unimpeded around and between neutral networks. The communities that interact in Cyberspace draw global membership because the Internet consists of networks that span the globe: underlying these networks are physical cables, routers, radio waves and so on. Although governed by code, the Internet is ultimately a physical entity. Far from extraterritoriality, this physical infrastructure and its users are at all times physically within some national jurisdiction, and thus potentially subject to legitimate physical interference by the governing state. This is the ‘cyberlaw fallacy’, which discredited the cyberlibertarian school.

Carolina (2008) argues that because the physical layer of the Internet grounds Cyberspace within state boundaries, it is nonsensical to describe the Internet as ‘borderless’. The Internet has entered an “age of de-globalisation”: users’ physical locations are approximable and increasingly used to localise the content displayed to them. Further, Carolina (2011) warns that states may instruct service providers to ‘nationalise’ the Internet, severing connections to foreign nations and creating a network ‘island’. The recent troubles in Egypt, Tunisia and Libya led to intentional blocking and ‘Internet blackouts’ in attempts to chill discourse and impede organised revolt.\(^\text{12}\)

Clearly, even brief ‘Internet blackouts’ carry significant costs domestically and internationally. Although such action is unlikely to be sanctioned in democratic states, the recent blackouts are a reminder that the Internet is not an invincible medium. Despite the dominance of First Amendment values online (Murray 2010a), even a fully global network is not necessarily a liberal forum: the Great Firewall of China is a famous example of state censorship and action against

\(^{12}\)Libyan hosts were unreachable at the time of writing (spring 2011) and considerable periods of outbound inactivity were reflected in the Google Transparency Report.
cyberdissidence, while North Korea forbids access to the Internet entirely, operating a closed *intranet* in its place. Censorship sets a “dubious moral precedent” (McGuire 2007), setting into sharp relief the moral virtues of network neutrality.

Given the potential for physical and logical interference with the network of networks, liberal nations must be careful not overreact to cybercrime. Even in liberal jurisdictions, the ‘war on terror’ has arguably led to a trade-off between liberty and security (Waldron 2003): although many cyberharms, including cyberterrorism, may more easily be mitigated by network centralisation, to do so would be an authoritarian perversion of *lex informatica*, chilling discourse by sacrificing the open and neutral character of the Internet.

Morrison’s (2009) vicious attack on Barlow’s ‘Declaration of Independence’ concludes with a desire “to build and maintain the better, more humane cyberspace” that is the aim of the Declaration with “[a firm] imbrication in real-world politics”. Murray (2010b) seeks to achieve this by proposing a Bill of Rights for the Internet, criticising two earlier Internet rights projects. Common to all three projects are commitments to, *inter alia*: universal access; free expression; participation in the digital public sphere; data security; and network neutrality. Such projects could form the basis of international consensus to protect the network against interference both state and private: a constitutional framework for international co-operation on Internet policing, within a global digital ‘Schengen Area’ guaranteeing the free movement of bits.

*The Web of Trust*

Perfect regulation of a global, decentralised network is impossible, but if the regulatory settlement respects the norms and values of the active matrix, and the sophistication of *lex informatica* is such that subversion is costly to users, users are more likely voluntarily to comply with efforts to regulate. In the following chapter it will be argued that the codification of these norms ought to be undertaken or encouraged by state police forces, which should endeavour actively to participate in the active matrix.
“The Web is needed to support all sorts of relationships, on all levels, from the personal, through groups of all sizes, to the global population” (Berners-Lee 2000). Across a medium so pervaded by pseudonymity and the risk of computer-mediated harm such as identity theft, crucial to these relationships – whether personal, economic, or professional – is trust. Offline, trust-enhancing and risk-reducing operations such as secure transit, identity verification systems, consumer watchdogs and product review networks are commonplace. Similar systems exist online, serving to enhance the ‘Web of Trust’ (Berners-Lee 2000). Though this term is generally used to refer to trust-enhancing code, it can usefully be applied to contributions from other modalities: the eBay Feedback System is fuelled by users, markets and community norms, while others such as Symantec VeriSign and McAfee SiteAdvisor apply (for a price) trusted ‘seals of approval’ to centrally-approved websites.

Though state police departments have no business issuing product recommendations, one way in which police forces might enhance the Web of Trust is to offer cybercrime prevention advice to individuals and businesses. Simple measures such as ensuring users are aware of the risks posed by the Internet, are introduced to trust-based Internet browsing, are running up-to-date security software, have secured any wireless networks and are careful to monitor their children’s use of technology could complement existing physical crime prevention advice, thereby improving the utility of Cyberspace and communicating police attention to cybercriminality, participating in and strengthening the active matrix whilst securing public awareness of and confidence in Internet policing.

3.3 The nature of ‘Cybercrime’
It is not the intention of this section to provide a comprehensive overview of the many, varied and constantly developing range of cybercrimes. Rather, this

13 Notably, ‘Web of Trust’ is a cryptographic term of art, referring to a decentralised key exchange system.
section complements the preceding discussion of the challenges of regulation in Cyberspace with a sense of the breadth of harmful activities online.

Cybercrime is a broad and emotive term with roots in popular culture. So many harmful activities are today mediated in some way by technology that the term inspires a “new geometry of harm” (McGuire 2007). The National e-Crime Strategy (ACPO 2009) is built around a ‘Harm Framework’ spanning six dimensions: physical, social, environmental, economic, structural, and reputational. The near-ubiquitous spread of technology has led to the increasing fragmentation of communication across different networks (Home Office 2009), creating unprecedented digital forensic opportunities in the detection of all crime, including transmission logs and cell site analysis. The recent case of Kennedy¹⁴ found that the Regulation of Investigatory Powers Act,¹⁵ which balances the right to privacy against the public interest in criminal investigation, is compliant with the European Convention on Human Rights, confirming that digital forensics will continue to feature in the detective’s toolkit for the foreseeable future.

Three basic categories of cybercrime are identified by Wall (2007): computer integrity crimes, computer-assisted crimes, and computer content crimes. With notable exceptions, many cyberharms are de minimis: though it is not for the police to determine the severity of harms, their true seriousness often lies more in their aggregate impact than their marginal effect (Wall 2007). Computers and the Internet enable and restrain crimes both old and new, ranging from fraud and obscenity to digital trespass and virtual theft (Murray 2010a; Hanstock 2009 on the ‘right virtuālis’). A considerable market for countermeasures such as virus scanners and backup solutions has emerged, to the point that the weakest points in many ‘secure’ computer systems are users themselves (Davies 2000).

Wall (2007) describes three generations in the development of cybercrime, relabelled for present purposes. The first generation saw computers used as e-

---

jemmies: crimes were committed within discrete computer systems, to facilitate or prepare for plainly criminal acts such as theft, fraud, terrorism and so on. The second generation, network crime, is perhaps the most significant, as computer hackers exploited the connectedness of modern computers to inflict harms remotely. The third generation of cybercrime can be called distributed crime, mediated wholly by technology, using networks to distribute and automate the infliction of harm: distributed denial of service attacks use drone computers recruited subversively (by hackers) or politically (by hacktivists) into remotely-controlled ‘botnets’ programmed to flood target networks to the point of overload, rendering services inaccessible to legitimate users. Such threats are sui generis, “solely the product of opportunities created by the [I]nternet”, ranging as widely as virtual ‘rape’ (Dibbell 1993), cyberterrorism and cyberwarfare.

Wall predicts that as networks become increasingly responsive to human presence and interaction, a fourth generation of cybercrime will emerge. Wright et al. (2008) warn of the many “threats and vulnerabilities facing our privacy, identity, trust, security and inclusion in the rapidly approaching world of ambient intelligence” as we advance toward the ‘Internet of things’ (see von Kranenburg 2008). As the technologies behind ambient intelligence approach market readiness, regulators must be sure that closely researched safeguards are in place in order to defend against new criminal opportunities.

The aforementioned e-Crime Strategy defines cybercrime as “[t]he use of networked computers or Internet technology to commit or facilitate the commission of crime”. This is a clumsy definition, which while avoiding limiting its scope to harms committed across the Internet, draws an unnecessary distinction between ‘networked computers’ and ‘Internet technology’, and seems to be rooted in a fixed notion of criminality. The Harm Framework attached to the Strategy offers a more comprehensive vision of ‘criminal’ consequences in a networked environment, but the link between this and the narrower word “crime” is not clear on the face of the report. Criminality is an adaptive concept that must be prepared to evolve with the technologies that provide criminal opportunities – and the talents of those who identify and exploit them.
This evolution must be smoother than it has been thus far. It took the acquittal of teenager David Lennon\textsuperscript{16} to highlight the need for the criminalisation of \textit{sui generis} distributed denial of service attacks: sections 33 to 36 of the subsequent Police and Justice Act 2006\textsuperscript{17} amended the Computer Misuse Act 1990\textsuperscript{18} to include this new harm, as well as related crimes, including selling access to a botnet. There is, it is submitted, a role for the police in keeping abreast of such developments, working with legislators to ensure laws capture new harms consistently between nations.

Investigative practices must similarly evolve: partnerships between police and security researchers are paramount. One of the most topical challenges presented by the Internet is linking the digital footprints captured in server logs to the individuals responsible for harmful behaviour. The functioning of the transport set is such that every computer has its own unique Internet Protocol address. However, these addresses are transient, and as Network Address Translation is commonplace, one IP address may in fact represent a single point of connection for an entire network of computers, for example in a home, business, or Internet café. Accordingly, the details of the customer held by the service provider may not in fact be those of the user responsible for the suspect transaction. User accounts may be shared, passwords may be compromised, and networks may be used without authorisation, especially wireless networks (BBC News 2009b). Similar problems face vehicle crime: vehicles are identifiable through their registration or chassis number, but may be driven by someone other than the registered keeper. Though compulsory registration, Automatic Number Plate Recognition and cell site analysis can provide geographical locations which may help to narrow down the user of an unregistered mobile phone or identify a driver, but vehicle crime investigators enjoy a potent tool: section 172 Road Traffic Act 1988\textsuperscript{19} imposes a liability upon registered keepers to identify the driver of a vehicle upon police request. Were a similar liability

\textsuperscript{17} Police and Justice Act 2006 c 48.  
\textsuperscript{18} Computer Misuse Act 1990 c 18.  
\textsuperscript{19} Road Traffic Act 1988 c 52.
imposed upon the customers of service providers however, it might not always be possible to meet this obligation, especially in the case of unauthorised network access or particularly large or open-access public networks (BBC News 2009a). Large businesses may be able to examine their own logs of user-level traffic, however private individuals are unlikely to possess the requisite technical literacy to render this an effective option. ACS:Law has been ridiculed in court for overstating the reliability of IP address evidence (Media CAT v Adams 2011;20 Which? 2009): police forces must be careful not to make the same mistake.

Partnerships with security researchers will advance both the investigation and prevention of crime. However, crimes cannot be governed by technology: only detected, and in some cases prevented, by code. Governance of crime implies something far broader than mere target hardening and detection, including the prosecution, conviction, rehabilitation and diversion of e-deviants. To claim that technology is the solution to cybercrime is to ignore its human cause. Vehicle thefts are not a design problem: similarly, leaving one’s front door open might annul an insurance claim but is not a defence to burglary. The Internet is a social creation, thus the harms it mediates and enables are ultimately a product of that society. Though the ‘deviant intelligentsia’ may regard the subversion of security mechanisms as a technical challenge, the deeper issue is normative. The police should engage with these talented coders, and encourage them to contribute to crime science in a non-criminal capacity. Where this is not possible, robust extradition procedures are key.21

Similarly, social definitions of crime, and of harm, must be reflected in international legal definitions of the same. The ‘Twitter Joke Trial’ of Paul Chambers is a prime example of a justice system disconnected from the society it seeks to govern (Wainwright 2010; Murray 2010c), as is the case of a Chinese woman sentenced to a year in a labour camp over a retweet (Amnesty 2010).

---

20 Media CAT Ltd v Adams and others [2011] EWPCC 6, 8 February 2011.
21 At the time of publication, there is a concern that should the UK use Protocol 36 of the Lisbon Treaty to opt out of EU criminal law, our ability to police international crime—including cybercrime—will severely be hampered. See a forthcoming paper from Hinarejos, Spencer and Peers (2012), a link to a summary of which appears in the bibliography. The case of Gary McKinnon is an excellent example of the issues at stake in the extradition of cybercriminals.
Networked governance of crime accepts crime as a social, not a technical, problem, and is reactive to and constitutive of social norms. The following chapter offers a schematic for this governance.

Chapter 4 – Policing and the Digital Community

The previous chapters have provided overviews of the nature of policing and the substance of Cyberspace. This chapter brings the two together, suggesting areas for further research and proposing a loose schematic for ‘peace, order and good governance’ online.

4.1 Towards a networked model of policing
Writing on the distinction between ‘policing’ and ‘the police’, Crawford (2003) proposes four models by which police forces might adapt to an environment wherein policing functions are shared with other institutions. The force can (a) integrate policing partners into an extended police family; (b) steer policing partners from a distance; (c) establish a network of equals engaged in the provision of security; or (d) allow the market to structure the relationships between security providers. This chapter will apply this logic to the digital environment, arguing that the networked model is the only approach suited to the unique characteristics of the Internet.

It was shown in Chapter 3 that lex informatica is largely under the control of private Internet Information Gatekeepers, popular nodes within the network through which information is frequently stored and retrieved – household names such as Google and Facebook for example – and that responsible gatekeepers embrace certain human rights responsibilities (Laidlaw 2010). In shaping the discourse and interactions of the gated, gatekeepers perform a social function; and as code is supreme (Lessig 1999), their power in so doing is essentially unlimited. As members of a networked community, the gated will engage the active matrix and feed back to the gatekeeper their approval or disapproval of the changes, often via the market (Murray 2010a). However, the
market dominance of the largest gatekeepers may render this an impotent mechanism in all but the most extreme cases, offering little protection to disproportionately affected minorities: Cavanagh’s market model thus fails. Government intervention in market failure quickly runs into a problem on the open Internet: which government should legislate to intervene? A gatekeeper may act as an arbitrageur, seamlessly relocating its assets and servers to a lower-cost jurisdiction.

Arbitrage would not be an option if states, all committed to the open network, agreed to legislate on Internet matters with one voice. An international framework for the consistent application of these standards ought to be established. The beginnings of this consensus, it is submitted, are to be found in Murray’s Bill of Rights for the Internet (2010b). Wall (2007) points out that states will inevitably disagree on such matters, but this only strengthens the argument to begin the process sooner rather than later: as Murray points out, international treaties do not form overnight. The existence of similar projects suggests there is appetite for a covenant between states to protect certain principles of the information age: Morrison’s (2009) vicious attack on Barlow’s ‘Declaration of Independence’ concludes with a desire “to build and maintain the better, more humane cyberspace” that is the aim of the Declaration with “[a firm] imbrication in real-world politics”.

This ‘imbrication’ is key. Regulators must approach this consensus not as ideological competitors but as nodes within a network striving towards a degree of consistent governance for an international agora. “We can view Cyberspace much like the United States [before 1791: a] virgin territory which we may shape to reflect our values, culture and society” (Murray 2010b). Cyberspace is neither a territory nor is it virginal: existing communities and regulatory networks must be included in the network, and it is false to presume that a unified cyber-*demos* has emerged between fragmented international microcommunities. Nonetheless, Cyberspace is “[a] global space, not a domestic space” (Murray 2010b), thus domestic regulators and police forces must globalise, recognising their place as
nodes within a global network of regulators, some with more regulatory gravity (Murray 2011) than others.

Covenants as to shared outcomes and consistent means would strengthen law as a modality of regulation in Cyberspace, supported by co-operative transnational detection, enforcement and information sharing along common procedures (Harfield 2011). Adjudication of this covenant could be carried out by the World Wide Web Consortium (W3C), famous for its technical standards and commitment to the openness and accessibility of the network, without compromising the democratic credentials of the covenant. The W3C could administer standards for good governance online, applicable to signatory states and the Internet Information Gatekeepers within their collective purview. International partnership between legislators, service providers, gatekeepers and user groups could thus create a safer, more consistent, network for all, with the interests of each party protected through discourse on the balancing of ‘rights’, encouraging responsible governance and developing more unified regulatory norms.

4.2 Penetrating the active matrix

Nyabuga (2011) advocates the deterrence of cybercrime through a binary partnership between enforcement agencies and service providers bent upon the identification and prosecution of cybercriminals. Enforcement alone is simply not the answer. In the previous chapter it was argued that effective regulation online relies upon a symbiosis between network communities and network governors, requiring legislators and gatekeepers to step into the ‘active matrix of pathetic dots’ in order to produce responsible, norm-compliant regulatory settlements which inspire popular voluntary compliance. Regulators must engage with the active matrix in order to identify and develop the relevant norms, working with gatekeepers and other digital architects to codify these norms into lex infomatica. Failure to do so will result in a deviant and alienated “intelligentsia” (Cavanagh 2007), particularly if they are themselves the focus of enforcement action (David 2010).
Understanding the Internet as a collection of fragmented microcommunities, rather than a unified public sphere (Jeffries 2010), enables a localisable vision of an otherwise global space. This fragmentation occurs as pathetic dots interact through specific gatekeepers, forming network communities and engaging in discourse on the terms set by gatekeepers who become intrinsic to the identity of the user, his digital footprint and even his self-perception, shaped by his patterns of interaction with other network nodes (Turkle 1995).

Chapter 2 explained that although there is only a weak link between policing and crime control, “uniformed patrol remains the bedrock of policing, and this will continue to be preoccupied with order maintenance, rather than crime control” (Reiner 2010). However, “the change from atoms to bits” (Negroponte 1995) has led to an interjurisdictional network of private spaces, within which users constantly ‘arrive’ but never seem to ‘travel’, leaving footprints without streets. Policing these private spaces is mostly carried out by ‘private security’ forces within self-regulating communities, patrolled not by public police officers but by private superusers and digital architects, setting the terms of the digital environment and enforcing behaviour in accordance with those terms. What, then, is the role of the public police?

Private gatekeeping leads to private justice, and we are fortunate that some of the larger gatekeepers, such as Google, have committed themselves to liberal values such as openness, privacy and transparency. Most remain ultimately committed to profitability, which Laidlaw (2010) argues should be addressed by assigning legal liabilities. Nonetheless, due to the decentralised nature of the Internet, the virtues of preserving network neutrality, and the supremacy of code in the online experience, many of the policing functions we entrust to the public police in real-space are inevitably concentrated in private hands. In order to ensure that public rather than private norms shape the inevitable discretion of private actors in applying their considerable powers, public police forces should monitor and work with private gatekeepers.
4.3 Safer e-Communities Teams

It is proposed that these aims could be achieved is by the expansion of specialist state ‘cybercops’, such as the Police Central e-Crime Unit. Inspired by the Safer Neighbourhoods initiative, whereby electoral wards are assigned their own dedicated police teams, Safer e-Communities Teams (SeCTs) could be assigned ‘sectors’ of the Internet, engaging in covert patrol (including ‘open source’ investigation or reconnaissance) and overt partnership with the gatekeepers within their ‘sectors’. For example, a social networking sector would take responsibility for Facebook, MySpace, Bebo and similar; a blogging sector would cover Blogger and WordPress; a shopping sector would cover eBay and major online retailers, and so on. Each would develop its own specialist knowledge and partnerships with other police agencies nationally and internationally, such as foreign SeCTs, fraud squads, counter-terrorism divisions and paedophile units, supporting investigations as a gatekeeper liaison.

SeCTs engaged in covert patrol of the Internet would harness the pseudonymity of the networked space, accessed through spoofed addresses with the assistance of service provider partners, identifying and reporting harmful and criminal activity for investigation and potentially prosecution. Overt patrol online is unlikely to be a possibility, due to the risks posed by pseudonymity and the likelihood of impersonation, even where gatekeepers are prepared to devise mechanisms for overt police presence in discussion forums and similar. The risks posed by decentralisation are simply too great.

SeCT patrol officers will not require police powers or police identification, except where crimes are identified and escalated for detection. For this reason, the patrol function could be carried out by police auxiliaries, perhaps volunteers within the ‘Big Society’, even working from abroad. A new, specialist ‘Police e-Community Support Officer’ could receive training on the identification and reporting of cyberharms, and could even ‘police’ communities of which they are already members.
Aside from patrol, SeCTs should take the initiative in establishing partnerships with gatekeepers, user groups and interest groups, encouraging public consultation on the governance of online communities. Some responsible gatekeepers already invite user input, such as Facebook Site Governance, or make rights-related decisions available to users, such as the Google Transparency Report. SeCTs should work with partners to ensure that the human rights impact of decisions implemented through code have been considered, respecting rights settlements across an international user base, perhaps issuing warnings where necessary to users identifiable as accessing services from countries where content might be offensive, avoiding the interjurisdictional fallout seen in the *Yahoo! v LICRA*22 proceedings.

The objective of SeCTs is not merely to maintain order, but to facilitate digital social life, providing moral leadership (*per* Sykes 1986) to commercial operators of *lex informatica*. An internationally agreed Internet Bill of Rights (*per* Murray 2010b) would structure and legitimise normative intervention supporting the ethic of responsible gatekeeping (Laidlaw 2010) while reducing or removing the incentive for arbitrage.

As well as encouraging user participation and rights-consciousness, an important element in the deterrence of cybercrime is target hardening, through code and education. Forming partnerships with software developers and security researchers, SeCTs could guide and catalyse innovation in security technologies, whilst raising awareness of cybersecurity issues and encouraging users to adopt security-enhancing working practices, similar to current crime prevention initiatives. Projects intended to enhance the safety of children online are already in place, ranging from the nationwide ‘FBI-SOS’ cyber-citizenship programme to a bespoke course delivered to local schools and parents by a PCSO in Abingdon, Oxfordshire (Cunningham 2011).

A similar initiative emerged under the Prevent strand of CONTEST, the Home Office counter-terrorism strategy (Metropolitan Police Service 2010a). Under

---

22 *Yahoo! Inc. v La Ligue Contre Le Racisme et L’antisemitisme* 433 F.3d 1199.
this scheme, Safer Neighbourhoods Teams across London visited internet cafés to warn patrons against accessing extremist material, displaying an “Internet Code of Conduct” on posters and screensavers, reminding users that the cafés will report suspicious Internet activity to the police. The scheme has since expanded into libraries and other public places in the Tower Hamlets area, and deserves attention at a national level as a model for overt Internet police patrol: though Cyberspace itself cannot overtly be patrolled, its public points of entry, plainly within the jurisdiction of one state, can be. Although most people access the Internet from private spaces such as homes and businesses, convergence has led to Internet-equipped mobile devices and widespread availability of wireless networks and similar technologies. Always-on Internet connections blur the notion of a single ‘point of entry’ for Cyberspace as almost all public spaces are equipped with Internet connectivity. Just as businesses require employees to conform to acceptable use agreements for Internet use, a public Internet Code of Conduct could serve as a mutual statement of expectation between the police and the public. This message, developed by SeCTs in consultation with the online public, could be delivered in a manner similar to the national Policing Pledge, publicised through advertising in physical and digital spaces. In partnership with gatekeepers and service providers, SeCTs could investigate users suspected of violating the code of conduct, perhaps working with Safer Neighbourhoods Teams to remind users of their obligations in Cyberspace and address any unauthorised use of their Internet connections.

SeCTs could engage with the public via social networking and microblogging sites around issues within their respective sectors, encouraging Safer Neighbourhoods officers to do the same around local issues within their wards, working with partners to reduce the risks of, and detect and prosecute, impersonation: the Twitter Verified Account feature, used to authenticate celebrities and businesses, provides a gatekeeper-issued stamp to a user account once it has been confirmed as originating from its purported source. Public-police engagement online has thus far developed haphazardly, ranging from individual officers and teams posting on Twitter, for example @hotelalpha9, @PCStanleyWMP, @SgtGaryWatts, @cheshirecopter and so on, often with
corresponding blogs. @CO11MetPolice is used to great effect by the MPS to communicate with protestors during public order events, gathering intelligence on changes to the marches and enabling tactical decisions to be made to contain flashmobs and impromptu disorder. Some forces hold regular online surgeries, such as the monthly Leicester Constabulary Roads Policing Unit Web Chats. SeCTs could play an important role in encouraging and standardising the police use of Twitter and similar services, while ensuring that officers using social networking websites off duty are aware of their risks and responsibilities in doing so.

One important issue that SeCTs should address with gatekeepers, service providers and private businesses is the under-reporting of cybercrimes. Anxieties about risks to personal data and the economic cost of devaluation within the Web of Trust constitute a commercial disincentive to report cybercrimes to the police. SeCTs could pledge not to disclose investigations to the press, or even agree not to charge in cases where prosecution would not serve the public interest, instead using such reports to gather intelligence to assist in other investigations.

Effective online policing demands operational consciousness of international partners. Transjurisdictional investigation and information sharing must become the order of the day, and is likely to lead to a rise in European Arrest Warrants, Interpol circulations, and extradition proceedings. Such practices are certainly not new, and are used to great effect in combating international organised crime. Operation Lath for instance saw the Police Central eCrime Unit make use of intelligence from foreign police forces and the financial sector, leading to nineteen arrests, shutting down a ‘money mule herd’ exploiting the ZeuS computer virus to steal £6mn from UK bank accounts over a three month period (Espiner 2010). A similar arrest was made the following month (Metropolitan Police Service 2010). Such investigations are exemplary of the networked knowledge-work which underlies successful transnational co-operation, and ought to be rolled out to more routine crimes.
Chapter 5 – The Context and Character of the Digital Police

In Chapter 2, the seven core findings of policing research identified by Reiner were distilled into four broader headings. This chapter provides an overview of the applicability of these findings to the environment mapped out in Chapter 3, and makes suggestions for further research.

Bits flow across borders on the open network: thus, borderless international cybercrime calls for a borderless networked conception of policing. Organised international crime already receives a considerable degree of international police co-operation, however, this co-operation must extend to lower level cybercrimes if policing on the Internet is to be effective. Controlling deviance online requires a cultural shift in the perception of what counts as criminality and a reassessment of the idea that online criminality is somehow not a police matter. It is the police role to respect and enhance self-regulation by injecting concern for the public interest into private order maintenance through close partnerships with private gatekeepers.

5.1 A New World Order?
The distinction between the social function of ‘policing’ and the institution of ‘the police’ is even more salient when one conceives of the Internet as a transjurisdictional social space. Police forces act together only in cases of serious organised crime, with low-level Internet disorder going unreported, undiscovered, ignored, and even unrecognised. This state of affairs led to a popular misconception of the Internet as a state-irregulable extraterritory, wherein private individuals set and enforce rules within their own domains, technically connected but socially separate from one another, with inevitable overlaps at the user level. Some communities, such as Wikipedia, have enjoyed enormous success in self-regulation, using code and norms to establish and maintain order (Goldspink 2009); others have reacted to harmful behaviour by establishing an autonomous democratic order (Dibbell 1996). However, as the pervasive mediation of life by technology has developed, and the proportion of discourse and interaction online that has concentrated in private hands has
expanded (Laidlaw 2010), the case for guarantees of public-minded and rights-conscious governance online has emerged.

Offline, police forces are the apparatus of states, entrusted with the monopoly on the lawful use of coercive power to achieve just ends. Subscribing to a view of the Internet as social space, it is tempting to think of gatekeepers and coders as miniature sovereigns, multinational corporations, arbitrageurs beyond the reach of disparate national laws. However, as Cavanagh (2007) and Wall (2007) remind us, the Internet is both a social and a technological phenomenon. Technologically, the Internet is not inherently global (Carolina 2008, 2011), and is in fact anchored well within the physical jurisdiction of nation-states. The Internet is unlikely fully to deglobalise and thus centralise, so private governance is inevitable in a neutral, decentralised network spanning state sovereignties. What is required then is a networked approach to governance, recognising a nodal symbiosis of users, service providers, gatekeepers, police forces, researchers, lawmakers, and policy analysts (Bernal 2010; Laidlaw 2010; Murray 2010a).

Horizontal governance networks between states are advocated by Slaughter (2004) as the means by which the harmonisation of policy and practice is achieved, in order effectively to tackle global governance issues such as terrorism and serious crime and the challenges posed by interdependent economies and ecosystems. It is artificial to separate the international reach of communications networks from the globalisation of economies and crime, thus this dissertation essentially makes the case for Slaughter’s “New World Order” in respect of Internet governance. However, if policing on the network of networks is to be effective, deliberation must span far wider than state apparatus alone, involving user groups and private regulators at every stage whilst maintaining the values of the neutral and accessible Internet.

5.2 Patrol, Jim, but not as we know it...

Popular support for this regulation depends upon a policing model which prioritises engagement over enforcement. Regulators must recognise that their
first challenge is to dispel the perceptions of irregulability which are rooted in
the development of the Internet, inspiring appreciation of the new geometries of
harm online (McGuire 2007) and confidence in the means by which it is being
achieved. This demands a user-level normative shift, police inspiration of which
would have been dubbed “moral leadership” by Sykes (1986). Although just
coercion is expansive of human freedom, contrary to Nyabuga (2011), efforts to
police the Internet depend upon a constructive rather than an enforcement-led
approach: the Internet needs ‘peace officers’ before it needs ‘law officers’, at least
in respect of ‘low policing’ (Brodeur 1983), leaving remedies for low-level ‘order
maintenance’ largely in private hands, enabled by code and restrained by police
scrutiny and a rights-conscious ethic of responsibility.

Community-focused policing takes a variety of forms, but is principally
concerned with uniformed reassurance patrol offline. A pseudonymous,
decentralised, code-governed Internet however offers little scope for uniformed
patrol, with the exception of online surgeries and gatekeeper-supported
microblogging. The decentralised, end-to-end architecture of the network
renders Internet ‘stops and searches’ and ‘arrests’ undesirable and nonsensical,
and code places order maintenance functions almost entirely within the hands of
private superusers, moderators and administrators. Community policing online
thus heavily emphasises partnership working, ensuring that private gatekeepers
are maintaining order fairly and proportionately, and that overt criminality,
suspicious behaviour or information suggesting risk to children or other
vulnerable groups is immediately reported to the police, who then share it with
the relevant authorities in the relevant nations: Internet policing is thus very
much a “knowledge-work role” (Ericson and Hagerty 1997).

This monitoring can be achieved both through direct referrals from service
providers and covert Internet patrol, especially where gatekeepers are reluctant
to engage with police forces. Whether carried out by PeCSOs, police volunteers
or independent members of the community, covert patrol and commercially-
sensitive reporting is the ‘dragnet’ which classifies and refers problematic
behaviour to neighbourhood police officers familiar with technological issues or
specialist e-crime detectives, all attested officers who then have access to the full range of police powers and resources in assessing and reacting to the information received.

These partnerships and actions should be constituted by an international, mutual Internet Policing Pledge, accompanied by an Internet Bill of Rights. The charter should reflect the norms and concerns of existing self-regulating communities, gatekeepers, service providers and so on, underpinned by information-sharing agreements and common protocols between state police forces. The objective and mission of these police forces must be recast, not as merely order maintenance but as the facilitation of social life, through partnership working and norm-evaluation toward strong, self-regulating, public-minded digital communities, monitoring the local use of formal and informal sanctions while developing and promoting community-enhancing practices and technologies.

It was noted in Chapter 2 that the measure of effectiveness of the modern professional police appears to be public confidence rather than crime control. This is likely to remain so online, however, it has been shown that "uniformed governance" (Edwards 2009) is impractical in Cyberspace. Public confidence in digital policing is therefore likely to be measurable by the level of trust within digital communities, as the police work to facilitate the Web of Trust, through innovation with partners, effective detection, and the reinforcement of pro-social norms: the notion of a ‘right to security’ derivative from liability for failure to reassure (Ramsay 2008) could extend online into a ‘right to trust’, whether de facto or explicitly incorporated in the Internet Bill of Rights. Trust is, however, inevitably a measure of public confidence in ‘policing’, not in ‘the Police’: perhaps the latter will be determined by the attitude of the press, whose publications will influence popular opinion (see Reiner 2010). Success stories such as Operation Lath should thus be widely disseminated, reacting to negative media coverage by engaging with the active matrix via blogging, microblogging and social networking tools, participating openly in the marketplace of speech.
5.3 Discretion: the better part of justice?

It was shown in Chapter 2 that police work is highly situational, and the police thus enjoy – or are cursed by – a broad discretion. This is less likely to be problematic online: though much of this discretion is handed over to private actors, there is likely to be less irreversible immediacy online than in physical encounters. Certainly, instantaneous communication can create time-sensitive risks, and denial of service attacks can strike at any time; however, it is anticipated that instantaneous harms are more likely to be obviously criminal and thus deserving of swift disconnection and blocking by service providers, followed by investigation. Once matters become a matter for investigation, even offline, decisions are considered, logged and deliberated, supervised and monitored against standards, in a way that encounters on the street are not.

As actions online are readily recordable, police managers are more readily able to audit police action, including in covert patrol. If a patrol comes across a website which is likely to incite racial hatred for example, yet does nothing to report it, this will be identifiable through system logs and patrollers can be called to account for their inaction. This may lead to the phenomenon of digital ‘back-covering’, a popular complaint from officers whose actions have failed to stand up to scrutiny, thus further minimising the scope for prejudice in police decision-making.

As a great deal of the order maintenance function will remain in private hands, it is important to encourage gatekeepers to embrace the ethic of responsible gatekeeping (Laidlaw 2010). Private police action and the exercise of discretion must be rights-conscious, standing up to police and public scrutiny. In the offline world, private security associations such as the Security Industry Authority work in partnership with the police, receiving accredited training in the proportionate and legal use of force: similar training and accreditation could be administered by SeCTs.

It is worth noting that police culture studies (e.g. Loftus 2009) reveal an attitude unlikely to be receptive to networked governance of ‘low’ policing (Brodeur
1983). Although officers are increasingly technically literate, specialisation as a 
PeCSO will likely attract a new breed of ‘network police officer’, one who shares 
the commitment to policing within a network, who is articulate and creative 
enough to engage with the “intelligentsia” (Cavanagh 2007), and who is prepared 
to work with the private and public sectors in taking responsibility for making 
the Internet a safer, more cohesive social space. The aesthetic and situational 
qualities of police discretion will shape decision-making far less than in the 
world of atoms, and police isolation and conservatism will find little home in 
such a progressive and integrated new environment.

5.4 Facilitating fair social order, minimising risk online
Participation in a governance network necessitates acceptance that individually 
the police are marginal to social order, but collective action can shape behaviour 
and promote self-regulation. For instance, World Wide Web Consortium 
accessibility standards existed well before the Equality Act 201023 made website 
accessibility compulsory. Within a truly networked governance model, these 
existing international community standards would have been built into the 
legislative framework, adding weight to the pro-social community norm that 
websites ought to be accessible whilst empowering the international digital 
community to shape the quality of website accessibility. Instead, disparate 
standards will emerge based on commercial interpretations of the minimum 
standards for compliance with the ‘reasonable adjustments’ standard, 
compromising the uniform accessibility that was the objective of the Act. 
Coupled with “patchy at best” enforcement in the UK (McLelland 2010), one 
wonders whether such island legislation will ever stand for more than a 
toothless token nod to digital equality.

The role of ‘police advocacy’ is to identify opportunities for networked 
governance, allowing communities to set the standards through incumbent 
institutions with existing international normative weight. Policing must respect 
and trust the norms of the environment in seeking to enhance them, and thus 
inspire public confidence in Internet policing and governance generally. The

23 Equality Act 2010 c 15.
Police on this conception are an active entity, not merely an apparatus of the state, but “streetcorner politicians” (Muir 1977) representative of the global digital order. Police forces must of course remain concerned with detection, prevention & risk management, staying ahead of technological developments while adopting innovative new honeytraps and other investigative tactics.\textsuperscript{24} However, focusing on ‘safety’ alone is not enough. Rather, a new focus on inclusivity and the promotion of responsible community participation must emerge, thus attacking the structural causes of crime, disillusionment, alienation, and anti-social norms, bolstering the “internal policeman” (Durkheim 1893) crucial to self-regulation through education and inclusion in the network.

5.5 Final thoughts
This section organises some concluding remarks under each of ‘the five Ps’ that structure decision-making in the Metropolitan Police Service (Metropolitan Police Authority 2009), concluding with conjecture on the character of the road to networked policing.

**Presence:** As has been established, overt presence online is limited to microblogging, advertising and interactive surgeries. Though offering more ‘presence’ than the cardboard cut-out officer in some supermarkets, “uniformed governance” (Edwards 2009) online is simply not an option. Instead, covert patrol and partnership working will dominate the decentralised network, enhancing policy and public perception of Internet governance by police participation in the active matrix and improving the Web of Trust. The Internet Bill of Rights has the potential to become a high-profile project capable of capturing the public imagination and stimulating debate on cybercrime, raising the e-police profile.

**Performance:** The dissemination of success stories via the media and self-publication is important to establish an Internet presence. Crucial to successful enforcement is courtroom reception of evidential challenges such as IP address

\textsuperscript{24} For example, one of the key challenges for the police is to determine how to ‘seize’ domains and thus provide a visible deterrent (Saint 2010) in an age of caching, archiving, and (self-)republication, avoiding ‘the Fitwatch effect’ discussed in Chapter 3.
evidence, the pending appeal against the conviction of Paul Chambers, and the Wikileaks affair.

**Productivity:** The Police Central eCrime Unit is cited as saving £21 in potential theft for every £1 it costs to run (Downing 2011). With the possibilities of automated, volunteer and outsourced labour, these high returns should survive the expansion of focus advocated in this dissertation.

**Professionalism and Pride:** E-crime is a specialist area of policing deserving of high status internally and externally. Success through expansion requires good, competent leaders with a real sense of mission, and could benefit enormously from the recent recommendations from Neyroud (2011) intended to enhance the professionalism of police forces nationally and internationally.

Police culture on the Internet frontline will be very different from observational studies such as Loftus’ (2009), focusing on response and patrol officers. The duty to protect the vulnerable will become a duty to protect, inform and empower the technically illiterate. The distinction between ‘social work’ and ‘proper police work’ will vanish in a nodal conception of Internet policing, with partnership working spelling an end to ‘them and us’ conservativism. Aesthetic suspicion may disappear in pseudonymity or become more sophisticated as digital footprints become ever richer: observational studies of Internet police officers should pay close attention to this issue.

It has been established that although the digital social order is far from inoperable, tech-savvy ‘low’ policing is desirable in an increasingly tech-savvy world. The police have a role in working with self-regulating private webmasters to ensure that the norms which govern their actions are consistent with the responsibilities inherent in their exercise of the considerable power of code. The borderless network requires international co-operation if this is to be achieved. The police must not lose sight of robust responses to serious crime online, but be prepared to pass on some of the lessons learned from large transnational investigations to lower level cyberharms too. Else, cybercrime is in danger of
becoming the new ‘white collar’ crime: dealt with by private actors behind closed
doors, overlooked and underenforced, incubating antisocial norms and risking
norm creep and escalation. Police forces must work in partnership to address
disincentives to reporting and thus build a fuller picture of digital offending
(Mitchell 2008). Liability for failure to report, however, should be a measure of
last resort.

If the SeCT project is successful, even without overt routine patrol there are
dangers of looking back after a few decades and again assuming that policing is
crucial to the digital order. The police remain facilitators of social order, except
perhaps in the realm of serious and organised international crime, and this ought
to be made clear, so as to ward against centralisation and deglobalisation as
Carolina has predicted (2008, 2011). It is crucial to entrench neutrality and
openness in the networked policing constitution, protected the utility of the
Internet and its liberal character. Though policing at times of cyberseige may
become very different if network ‘islands’ are created as a countermeasure, any
subsequent investigation, defence and retaliation will be an intrinsically
international affair.

The chartered policing body proposed by Neyroud (2010) could play a key role
in advancing toward the requisite international consensus on the constitution of
networked policing for the network of networks. There can be no crime without
law, and Rome was not built in a day: thus, the sooner this ambitious project
begins, the better it will become. International forums on law enforcement
already exist, such as POLCYB and the Virtual Global Task Force, and could serve
as the foundations upon which the governance network is built. These proposals
are being made at a time of the following social developments. Firstly, the
information age continues to expand rapidly, and is increasingly global, with
cloud computing scattering single services across several nations. Secondly,
austerity measures in the wake of global recession, in particular the Big Society
initiative, seek to empower and strengthen communities and promote self-
regulation even in offline communities. Thirdly, and connected to this, renewed
faith in informal sanction and principles of restorative and rehabilitative justice
look set to help to empower said communities. The values advanced in this report are compatible with each of these developments, which could act as a catalyst to begin the process of drafting an Internet Bill of Rights: with the death of bin Laden arguably heralding the end of the war on terror, such a project may find new momentum in post-emergency states, and the threat to network neutrality may diminish.

Hot on the political agenda in the wake of Stuxnet and other acts of cyberespionage are calls from both sides of the Atlantic, and from Russia, for a “Geneva Convention in Cyberspace” (Rooney 2011). This is the perfect time to broaden the assessment of cyberharms into civilian cybercriminality: concerns over cyberwarfare could catalyse discussion of the roles and responsibilities of governments and police forces across the world in tackling cyberdeviance. This is an exciting time in the development of Internet regulation, and the values protected today will determine the shape of tomorrow’s world.

RICHARD HANSTOCK
www.ecrimeblog.com

Bibliography


Table of Authorities

**Cases**

*Kennedy v United Kingdom (2011) 52 EHRR 4* ................................................................. 134

*Media CAT Ltd v Adams and others [2011] EWPC 6, 8 February 2011* ..................... 137

*Yahoo! Inc. v La Ligue Contre Le Racisme et L’antisemitism 433 F.3d 1199* ............ 143

**Statutes**

*CAN-SPAM Act of 2003 15 USC 7701* ........................................................................ 127

*Computer Misuse Act 1990 c 18* .................................................................................. 136

*Digital Economy Act 2010 c 24* ..................................................................................... 129

*Equality Act 2010 c 15* ................................................................................................ 151

*Police and Justice Act 2006 c 48* .................................................................................. 136

*Police Reform Act 2002 c 30* ......................................................................................... 116


*Road Traffic Act 1988 c 52* ........................................................................................ 136
Command Responsibility and Successor Liability

SARA FANTONI

1. Introduction

This article examines the application of the principle of command responsibility to the situation where a commander failed to punish his subordinates for crimes committed before he became their commander, hereinafter referred to as successor liability. The article focuses on the Hadžihasanović and the Orić decisions. The controversy lies in the different understanding of the relationship between the notion of customary law and the application of the principle of command responsibility. The article will conclude with considerations on the future of successor liability.

2. Command Responsibility

Command or superior responsibility is a form of liability for international crimes, not a type of crime itself. The fully-fledged principle was formulated for the first time in the aftermath of the Second World War. It was subsequently codified in Articles 86 and 87 of the Additional Protocol I and later incorporated in the statutes of all major international tribunals. According to the ICTY jurisprudence, three key elements must be met for command responsibility to attach: the effective command and control of the superior over his subordinate,

---

4 Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I), 8 June 1977.
the knowledge that the subordinates have committed or are about to commit crimes, and the failure of the superior to either prevent or punish the commission of crimes by the subordinates.6

3. Customary Law

Customary law is understood as the body of international law deriving from the practice of States accompanied by opinio juris, the State compliance with a practice considered an international obligation. In Nikolić, the AC clarified the methodology of application of customary international law, holding that “in absence of clarity in the Statute, Rules, and jurisprudence of the International Tribunal, the Appeals Chamber will seek guidance from national case law, where the issue at hand has often arisen, in order to determine State practice on the matter.”7

4. Successor Liability in Hadžihasanović

In the case of Hadžihasanović the AC majority established that, given the absence of clarity in the Statute, Rules and the absence of jurisprudence on the matter, it would need to analyse the customary law relating to successor liability.8 However, the minority disputed this approach holding, in accordance with Nikolić,9 that there was no need to refer to custom, since the principle of command responsibility should have to be first applied to the circumstances before deciding it was not sufficiently clear to cover the given case.10 Interestingly, the AC held in the same case (but on a different issue) that “where a principle can be shown to have been so established, it is not an objection to the application of the principle to a particular situation to say that the situation is

---

8 Supra at 1 ¶ 45.
9 Supra at 7.
10 Supra at 1.
new if it reasonably falls within the application of the principle.” Judge Hunt submitted that command responsibility does not impose a direct responsibility for the acts of the subordinate. In fact, in Halilović it was held that the superior-subordinate relationship does not require direct or formal subordination, but that the material ability to prevent or punish is sufficient for the purpose of establishing effective control. As for the mens rea, the superior cannot be presumed to have actual knowledge, but it may be established through circumstantial evidence. The actus reus of command responsibility lies in the omission to prevent or punish, two distinct legal duties for superiors. The existence of two separate duties applicable at different times has been widely recognised.

Ultimately, what is at the heart of the division between the majority and minority in Hadžihasanović, is the underlying rationale of command responsibility, i.e. had the superior officer exercised his command authority properly, the offences would not have occurred. This determines the controversial outcome that, while being charged for omission to prevent or punish the commission of the crime by his subordinates, the superior is held responsible not for his omission but for the actual offences committed by his subordinates.

5. Successor Liability after Hadžihasanović

After Hadžihasanović, the ICTY was again faced with the issue of successor liability in Orić. Curiously, Judge Shahabuddeen, dissenting in Hadžihasanović,
upheld the Hadžihasanović decision in Orić on the basis that he was bound to follow the established precedent despite disagreeing with it. Judge Liu submitted that the error in interpretation of Articles 86 and 87 of Additional Protocol I represented a cogent reason to reverse. However, the AC missed the unique chance to overturn its wrong holding in Hadžihasanović, creating two binding precedents from which it will be very hard to depart in the future. While command responsibility was introduced precisely to avoid that superiors could escape criminal responsibility after having simply buried their heads in the sand or delegated the “dirty work” to the men on the ground, it would arguably be unfair – to give an extreme example – to hold President Obama liable for the offences committed under Bush’s administration.

6. Conclusion

A third way could be, departing from the customary principle of command responsibility, to look at Article 28 ICC Statute as introducing the new requirement of a nexus of causation between the superior’s failure to prevent crimes and the commission of those crimes, which might lead to a new application of command responsibility by the ICC. Paired with the AC jurisprudence, this could lead to the interesting hypothesis of the future development of a separate new crime of failure to punish. It will be up to the International Criminal Court to decide whether to depart from the ICTY jurisprudence or whether to continue the trend of command responsibility as a mere mode of liability, with its related limitations and controversies.

SARA FANTONI

20 Supra at 2, Declaration of Judge Shahabuddeen ¶ 2.
21 Supra at 2, Dissenting Opinion of Judge Liu ¶ 28.
22 Supra at 2, Separate and Partially Dissenting Opinion of Judge Schomburg ¶ 29.
23 Supra at 15, p. 25.
26 Supra at 15, p. 21.
BIBLIOGRAPHY

BOOKS AND ARTICLES


CASE LAW


To what extent are Economic, Social and Cultural Rights applicable in times of Armed Conflict?

MICHEÁL T MURPHY

INTRODUCTION

International Human Rights law, since its post World War II inception, has had at the heart of its legal discourse the concept of Civil and Political rights, and usually this has been to the detriment of meaningful discussion on Economic, Social and Cultural rights. The reasons that render this situation in international Human Rights law are manifold and vary from a lack of political will to the ineffective complaints mechanisms contained within the 1966 International Covenant on Economic, Social and Cultural Rights. However, while there still may be an apparent neglect for substantive international recognition of Economic, Social and Cultural Rights, there is evidence that the international community is always going further to enhance the notion of indivisibility of Human Rights. It has been stressed by many, including the United Nations, that “the indivisibility and interdependence of human rights are facts of political life which can be observed by everyone dealing with human rights in situations of conflict”.

Knowing then that the position of Economic, Social and Cultural Rights is largely solidified in International Human Rights law, this essay seeks to address the question as to the extent of the applicability of Economic Social and Cultural Rights in times of armed conflict. In order to do this I intend to address firstly the

1 Thanks to Dr Jean Allain, Queens University, Belfast School of Law.
5 As enunciated by the United Nations in the Bangkok Declaration in which it “reiterated the interdependence and indivisibility of economic, social, cultural, civil and political rights”. http://law.hku.hk/lawgovtsociety/Bangkok%20Declaration.htm.
relationship between International Human Rights Law and the Laws of War, and to address some of the difficulties in applying these respective Legal Standards in times of armed conflict. I then intend to look at the application of the International Covenant on Economic Social and Cultural Rights in times of armed conflict, paying particular regard to the issue of derogations and extra-territoriality. Using the respective rights to water, healthcare and an adequate standard of living, I then wish look at the various obligations respectively imposed by each branch of the law in relation to Economic Social and Cultural Rights in times of armed conflict. I then intend to look at the recent report of the UN Fact Finding Mission in Gaza, to render a recent appraisal of the position of Economic, Social and Cultural rights in times of Armed Conflict.

"OH EAST IS EAST AND WEST IS WEST AND NE'ER THE TWAIN SHALL MEET?"7 – THE RELATIONSHIP BETWEEN INTERNATIONAL HUMAN RIGHTS LAW AND INTERNATIONAL HUMANITARIAN LAW

International Human Rights Law and International Humanitarian Law have developed in a rather contrapuntal fashion, two separate areas of the law with two completely separate mandates, one to govern war and the other to operate in times of peace, yet both having equal significance in international legal order.

The express purpose of the Laws of War or International Humanitarian Law is to “regulate the relations between nations by providing rules for the protection of certain categories of enemy persons in times of war”.8 While the laws that govern conflict are some of the oldest in the international legal framework9 the phrase “International Humanitarian Law” is relatively recent.10 It is perhaps of interest to note that the development of International Humanitarian Law is inextricably linked with major atrocities. The American Civil War introduced the Lieber Code, The Battle of Solferino bred the Red Cross movement and the atrocities of World

---

7 Rudyard Kipling, The Ballad of East and West.
8 Schindler, D Human Rights and Humanitarian Law American University Law review 1982 (31) p935. Some early examples of this can be found in General Order 100 Instructions for the government of armies of the United States in the field (The Lieber Code).
9 ibid.
10 ibid.
War II introduced the Geneva Conventions and Genocide Convention, which led to a development of the notion of state, as opposed to individual, criminal responsibility.\textsuperscript{11}

Human Rights Law, which had established itself in various different constitutional frameworks of nation states,\textsuperscript{12} did not appear in the international arena to the end of World War II.\textsuperscript{13} It was the atrocities committed during that war that led to the implementation of the Universal Declaration of Human Rights, which has now largely become part of customary international law and began the creation of International Human Rights Law framework.

In spite of the fact that \textit{prima facie} these two arms of international law share a commonality in attempting to bring about a sense of regulation to the atrocities that are committed in war or the atrocities that are committed in peace time,\textsuperscript{14} they have until recently remained completely separate areas of law. This however was to change in the 1960s as the international community came to recognize, not only the parallel that exists but the importance of merging these two aspects of the law in order to ensure foremost that in the regulation of international law, whether in peacetime or times of armed conflict, was the notion of humanity. It was a process that Meron quite rightly refers to as the \textit{Humanization of Humanitarian law}.

In response to a number of armed conflicts that had erupted in many parts of the world in the 1960s,\textsuperscript{15} the UN met at the Tehran Peace conference in 1968, and adopted GA Resolution 2444,\textsuperscript{16} which emphasized the necessity of “applying basic humanitarian principle” as a means of respecting Human Rights in times of

\textsuperscript{11} Meron, T \textit{The Humanization of Humanitarian Law} American Journal of International Law 2000 (94) p243.
\textsuperscript{12} As demonstrated by The United States Declaration of Independence (1776) and The French Declaration Rights of Man and the Citizen (1789).
\textsuperscript{13} Schindler, D [\textit{supra note 7}] at p935.
\textsuperscript{14} Meron, T See Generally [\textit{supra note 10}].
\textsuperscript{15} In Particular Vietnam, Nigeria-Biafra and India Pakistan, for others see \textit{note 7} at pp 937.
armed conflict.\textsuperscript{17} The Symbiosis of International Human Rights law and International Humanitarian Law is evidenced further when we look at how some of the key instruments of the laws of war contain express provisions that are indicative of Human Rights language. The Martens Clause\textsuperscript{18} contained in the Second Hague Convention of the Laws and Customs of War on Land is a clear indication that even in times of war the highest standards of civility and conscience must be sought in order to ensure the proper adherence to the ‘principles of humanity’. There is no doubt that the Martens Clause is an indication of how the laws of armed conflict can interact with Human Rights Law and “is arguably an expression of what they both have in common”.\textsuperscript{19} As noble in language as the Martens clause may be, there is no doubt that it does prove an inadequate legal formation. Meron argues that the ethical and humane language adapted is compensatory for a lacklustre legal framework and it represents a pull towards normativity.\textsuperscript{20} Article 75 of Additional Protocol 1 to the 1949 Geneva Conventions is also a framework that is notable for its Human Rights language. It affirms that victims of an armed conflict shall be afforded the ‘fundamental guarantee’ of protection based on non-discrimination and protection from “outrages upon personal dignity, in particular humiliating and degrading treatment, enforced prostitution and any form of indecent assault”.\textsuperscript{21} The same can also be said for Article 6 of Additional Protocol II of the Geneva Conventions of 1949, which states that principles correlative to Human Rights law\textsuperscript{22} are also applicable in International Humanitarian law specifically for the protection of victims of non-international armed conflict.

\textsuperscript{17} This was also emphasized by the United Nations in GA Resolutions 2652 (3 Dec 1970) 2674, 2678 (9 Dec 1970) and 2707 (14 Dec 1970).
\textsuperscript{18} “Until a more complete code of the laws of war is issued, the High Contracting Parties think it right to declare that in cases not included in the Regulations adopted by them, populations and belligerents remain under the protection and empire of the principles of international law, as they result from the usages established between civilized nations, from the laws of humanity and the requirements of the public conscience.” Second Hague Convention of 1899 and a variant of the same clause also appears in the IV Hague Convention of 1907.
\textsuperscript{20} Meron, T [Supra Note 10] at 245.
\textsuperscript{21} Additional Protocol 1 of the 1949 Geneva Conventions Art 75 2 (b).
\textsuperscript{22} In this instance it is issues surrounding procedural fairness and fair trial provisions, the presumption of innocence and the abolition of the death penalty; see Article 6 Additional to Geneva Protocols of 1949 at http://www2.ohchr.org/english/law/protocol2.htm
International treaty bodies have also played an important role in facilitating the amalgamation of Human Rights and International Humanitarian law. In the case of Abella v Argentina\(^{23}\) the Inter-American Commission on Human Rights stated that "The American Convention, as well as other universal and regional human rights instruments, and the 1949 Geneva Conventions, share a common nucleus of non-derogable rights and a common purpose of protecting human life and dignity."\(^{24}\) The European Court of Human Rights in the relatively recent decision of Kononov v Russia\(^ {25}\) has also been seen to apply standards of International Humanitarian Law in a Human Rights forum,\(^ {26}\) thereby accentuating the position that at a regional level, these two variants of international law can be compatible with each other.

In a rather ironic twist of fate one of the strongest cases for the applicability of Human Rights law in times of armed conflict, is actually the means by which those same rights may be temporarily suspended. Where a situation requires, i.e. a state of public emergency that threatens the life of a nation, a state party is free to derogate from some of its international human rights treaty obligations.\(^ {27}\) It is important to note that derogation from international treaty obligations does not steadfastly mean a departure from human rights standards generally, and those rights protected by international law as being non-derogable.\(^ {28}\) While much has been said on the issue of derogations in relation to armed conflict\(^ {29}\) which would tend to indicate that International Humanitarian Law becomes lex specialis, there is also a tendency to regard Human Rights standards to have precedence, as has

\(^{23}\) Case 11.137.
\(^{24}\) ibid at para. 158.
\(^{25}\) Application no. 36376/04.
\(^{26}\) This Was much to the Criticism of Pinzauti who stated "The ECHR, although formally applicable in times of armed conflict, is not designed to regulate such exceptional situations. Instead humanitarian law, which is the applicable lex specialis, is better tailored to regulate the belligerents' behavior on the battlefield" See Pinzauti, G The European Court of Human Rights Incidental Application of International Criminal Law and Humanitarian Law. International Journal of Criminal Justice (2008) 6 p1060.
\(^{27}\) International Convention on Civil and Political Rights (ICCPR) Art 4 (1). The European Convention of Human rights also contains such provisions in Article 15(1)
\(^{28}\) Jus Cogens or Peremptory Norms.
\(^{29}\) See ICJ Advisory opinion of The Wall Which states that "the protection offered by human rights does not cease in the case of armed conflict, save through the effect of provisions for derogation of the kind to be found in article 4 of the ICCPR" or more generally Gross, O and Ní Aoláin, A Law in times of Crisis: Emergency Powers in theory and Practice Cambridge University Press. 2006.
been noted by the UN General Assembly “fundamental human rights, as accepted in international law and laid down in international instruments, continue to apply fully in situations of armed conflict”.

It is apparent that these two branches of international law have been in a symbiotic yet adversarial relationship; never quite knowing when one ends and the other begins, and the response of the international legal community to this has been to create a semblance of unified law. This indeed is a far removed from the opinion of Draper who believed that the laws were diametrically opposing and that they should, as such, be kept in their respective spheres for fear of harming them both. Lubell rather quaintly points out to us the difference in the two arms of the law by explaining the dichotomy of language used in both how for example, a military officer finds the notion of a right to life in an armed conflict equally as vexing as a human rights lawyer would the notion of a military objective. However, as I intend to demonstrate, international legal thinking is so substantially developed on this particular issue that International Law is now prepared to give significant deference to the applicability of Human Rights in armed conflict and that this notion is as equally pertinent to the discussion of Economic, Social and Cultural Rights as it is to Civil and Political Rights.

THE APPLICATION OF THE INTERNATIONAL COVENANT ON ECONOMIC, SOCIAL AND CULTURAL RIGHTS (ICESCR) IN TIMES OF ARMED CONFLICT – DEROGATIONS AND EXTRA-TERRITORIALITY

As has already been confirmed by the United Nations, and in spite of the ambivalences between Human Rights law and Humanitarian Law, international human rights standards continue to apply during times of armed conflict. The

30 UN GA Resolution 2675 (XXV).
31 Draper, G Humanitarian Law and Human Rights Acta Juridica (193) 1979
33 See above at note 29.
34 A Position which is also confirmed by the International Law Commissions (ILC). See The effect of Armed Conflict on Treaties: An Examination of Practice and Doctrine: Memorandum by the Secretariat at http://untreaty.un.org/ilc/documentation/english/a_cn4_550.pdf. The Document also states observances by the International Law Institute that "the existence of an armed conflict
question that then arises is whether or not this application, of International Human Rights Standards extends to the International Covenant on Economic, Social and Cultural Rights (ICESCR). As has already been mentioned, the existence of derogations in international human rights law provides a strong case that human rights treaties continue to exist in times of armed conflict, even if they are partially suspended. However in the case of ICESCR, however, no such derogation provision exists. Mottershaw is keen to point out that the absence of a derogation clause in ICESCR does not provide an a priori reason that the treaty does not apply in times of armed conflict.35 In spite of a lack of a derogation clause, ICESCR does contain a limitation clause in Article 4.36 It is important to note the fundamental conceptual and legal difference between a limitation clause and derogation from a right. A limitation clause may be invoked generally by the Human Rights system so as to ensure that the Human Rights system remains workable and that states retain an ability to regulate conflicts of interest that may occur from time to time.37 Derogations must only be invoked by a state in exceptional circumstances where the life of the public is under threat or in danger.38 It has then been argued that there would need to be a rather creative interpretation of Article 4 in order to allow for a suspension of Economic Social and Cultural Rights in times of armed conflict, or else that "there is no clear allowance for derogation or restriction in such times."39

The International Court of Justice in its advisory opinion on The Wall has also stated that ICESCR applies in times of armed conflict and that Israel is bound to apply the covenant to all territories within its effective control.40 The Committee on Economic Social and Cultural Rights has also sanctified Economic, Social and

---

36 "the State may subject such rights only to such limitations as are determined by law only in so far as this may be compatible with the nature of these rights and solely for the purpose of promoting the general welfare in a democratic society."
38 Ibid.
40 ICJ Advisory opinion on The Wall para. 113.
Cultural Rights in times of armed conflict. General Comment 19 on the right to adequate housing and forced evictions has stated that even in times of armed conflict the right to adequate housing and not to be forced to eviction still applies.\textsuperscript{41} General Comment 15 on the right to water also asserts that in times of armed conflict the right to water must be upheld under obligations in International Humanitarian Law as well as ICESCR.\textsuperscript{42}

The notion of how Human Rights standards can apply extraterritorially is extremely important when addressing the question as to whether or not Economic Social and Cultural Rights apply in times of armed conflict. The Covenant on Civil and Political Rights has a rather broad extra territorial reach,\textsuperscript{43} where a state must extend protection not only to individuals within the state territory but also to individuals over whom the state exercises jurisdiction. Meron argues quite strongly that the breadth of this extraterritorial reach of the ICCPR is appropriate for any human rights instrument. “\textit{Narrow territorial interpretation of human rights treaties is anathema to the basic idea of human rights, which ensure that a state should respect human rights of persons over whom it exercises jurisdiction.}”\textsuperscript{44}

As is mentioned in customary international law, and made clear by the Vienna Law on Treaties, a state shall, unless otherwise stated or intended, be subject to a particular treaty only in respect of its own territory.\textsuperscript{45} In the area of territoriality however, ICESCR is once again lacking in any evident legal provision to determine on the application and scope of the treaty in times of armed conflict. It is also pointed out to us by Dennis, that the negotiating history of ICESCR would not indicate a trend of state willingness for the substantive obligations contained

\textsuperscript{41}http://www.unhchr.ch/tbs/doc.nsf/(symbol)/CESCR+General+Comment+7.En?OpenDocumen\textsuperscript{t} at para. 5.
\textsuperscript{42}http://www.unhchr.ch/tbs/doc.nsf/0/a5458d1d1bbd713fc1256cc400389e94/$FILE/G0340229.pdf at para. 21-22.
\textsuperscript{43}ICCPR Art 2 (1).
\textsuperscript{44}Meron, T \textit{Extraterritoriality of Human Rights Treaties} American Journal of International Law 89 (1) 1995 p82.
\textsuperscript{45}Article 29 Vienna Convention of the Law of Treaties.
in ICESCR to apply extraterritorially or in times of armed conflict.\(^{46}\) There was unwillingness on behalf of the Soviet Bloc states to adhere to territorial clauses, as they believed it would perpetuate the colonial system.\(^{47}\) However, there has been some notable international precedent established to affirm that ICESCR applies extraterritorially.

The Committee on Economic Social and Cultural rights has stated in relation to Israel that the covenant applies “to all areas where Israel maintains geographical, functional or personal jurisdiction.”\(^ {48}\) In a similar observation the Committee reinforced this position by stating that it ‘deplored’ the position that Israel took in stating that the covenant does not apply to areas that are within its sovereign jurisdiction and control.\(^{49}\)

The position of the International Court of Justice (ICJ) was also explicit in pointing out that the obligations contained within ICESCR are applicable extraterritorially. “...this [covenant] guarantees rights which are essentially territorial. However, it is not to be excluded that it applies both to territories over which a state party has sovereignty and to those over which that state exercises territorial jurisdiction.”\(^{50}\) Dennis has levied some harsh criticism at the court in this respect saying that virtually no analysis was provided to determine the reasoning behind the decisions. He states that “the negotiating history of article 4 [ICESCR]... strongly suggests that states should be free to derogate from their obligations under the ICESCR during armed conflict since the treaty [does] not contain provisions to the contrary.”\(^{51}\)

There are indeed many conceptual difficulties in looking at the application of Human Rights treaties extraterritorially and this is particularly pertinent to

\(^{46}\) Dennis, MJ Application of Human rights Treaties Extraterritorially in Times of Armed Conflict and Military Occupation American Journal of International Law 99 (1) 2005 p 128.

\(^{47}\) ibid.


\(^{50}\) ICJ Advisory Opinion on The Wall Para 106 and 112.

\(^{51}\) Dennis MJ [Supra note 45] at p 140.
ICESCR considering its notable silence on extraterritorial application. However, in spite of some academic and legal dissonance\(^{52}\) on this particular subject, I still believe that it is fair to say that human rights obligations may extend extraterritorially and operate in times of international and non-international armed conflict; and that this extension is as equally valid for Economic, Social and Cultural Rights as it is for Civil and Political Rights.

**WATER, HEALTHCARE, HOUSING – THE OBLIGATIONS IMPOSED BY ICESCR AND IHL IN TIMES OF ARMSD CONFLICT**

“The Normative frameworks of Human Rights law and Humanitarian law are diametrically opposed with respect to obligations... While Human Rights law imposes obligations mostly on the state rather than the individual, Humanitarian law seeks to directly regulate the conduct of individuals as much as that of the state.”\(^{53}\)

In order to understand just how Economic, Social and Cultural Rights can be engaged in times of armed conflict, this section will examine some of the obligations that are contained in ICESCR and Humanitarian Law regarding the protection of economic, social and cultural rights, it will seek to examine this through the prism of a right to water, a right to adequate housing and the right to healthcare in times of armed conflict.

\(^{52}\) For instance there is difficulty over the meaning as to what constitutes an effective control or what amounts to jurisdiction of a particular region as was seen in the debates following the European Court decision in the case of Bankovic v Belgium (52207/99). It has been noted above that the Committee on Economic Social and Cultural Rights has given significant regard to the application of ICESCR extraterritorially, however it has been stated that the committee has not always been prepared to outline states obligations, and “General comment 3 fails to clarify the scope of the of the covenant” as well this, the committees reporting guidelines have failed to indicate the scope of the covenant and have referred to rights applicable ‘in your country’ at risk of misrepresenting the scope. See Mottershaw, E [supra note 18] p 453. Also note the consideration that if Human Rights Standards are being applied extra-territorially in situations of conflict, then the content of those rights may be altered in light of principles contained within International Humanitarian Law. See Lubell, N [Supra note 31] p 740-741.

WATER

Article 2 (1) of ICESCR states that high contracting parties to the treaty will be obligated to use the maximum of its available resources and engage in international co-operation to ensure that the rights subscribed within the treaty are progressively realizable.\(^{54}\) Mottershaw suggests that the right to water provides us with a good example of a right upon which there is an immediate obligation to protect at a minimum standard in times of armed conflict.\(^{55}\) Much has been written on the recognition of water as a human right, and CESCR General Comment 15 imposes obligations upon states to respect protect and fulfil a right as well as imposing an immediate guarantee that the right is exercised without discrimination.\(^{56}\) It has been suggested that aspects of ICESCR that contain provisions relating to water are “core obligations... just as binding, immediate and of the equivalent level to those under the ICCPR.”\(^{57}\)

The Committee on Economic Social and Cultural rights has also stated in General Comment 3 that in situations where there is a strain on resources, or a recession in the economy or other factors\(^{58}\) of mitigation, the rights of the most vulnerable people in society must continue to be protected.\(^{59}\) It can therefore be ascertained that while the particular obligation to respect, protect and fulfil the right to water may be flexible, the right can still be engaged in times of armed conflict.

The concept of a right to water is also of importance in International Humanitarian law with the protection of water being afforded both implicitly and explicitly. Firstly Article 54 of Optional Protocol I of Geneva Conventions (1977) states that objects which are indispensable to the civilian population may not be made the subject of a target, this is inclusive of water facilities such as

---

\(^{54}\) With exception to some obligations which General Comment 3 suggests are deemed to have immediate effect.

\(^{55}\) Mottershaw, E [Supra Note 18].

\(^{56}\) UN Doc E/C.12/2002/11.

\(^{57}\) Cahill, A The Human Right to Water – a right of unique status: The Legal Status and normative content of the right to water” International Journal of Human Rights 2005 9 (3) p398.

\(^{58}\) While not expressly clear it may be interpreted that this provision extends to be inclusive of times of armed conflict [emphasis added].

\(^{59}\) CESCR General Comment 3 Para. 12 http://www.unhchr.ch/tbs/doc.nsf/0/94bdbaf59b43a424c12563ed0052b664?OpenDocument
installations, supplies and irrigation works.\textsuperscript{60} It is firstly a generally accepted principle of customary International Law that there is a ban on the use of poison, which would extend to the poisoning of wells and water supplies.\textsuperscript{61}

**HEALTHCARE**

The same may also be argued in relation to the protection of the right to health as proscribed in Article 12 ICESCR.\textsuperscript{62} Using very similar language to the obligations pertaining to the protection of the right to water, it is recognised by CESCR that there is an immediate obligation to protect the right to health.\textsuperscript{63} General Comment 14 also granted a fairly broad definition to the right to health to recognize the various social concerns against which such a right may be engaged. It is important to note that these concerns were inclusive of situations of global pandemics like the HIV/AIDS crisis and violence and armed conflict.\textsuperscript{64}

Even though the approach adopted by IHL in relation to the right to water is fairly clandestine, when it comes to the right to health, IHL is much more robust with regards to obligations and protections afforded. It can be argued that the entire premise of IHL is to afford protection to those who are engaged in or subject to acts of warfare and in need of healthcare. There are numerous provisions in IHL which protect the right to healthcare, the most authoritative being found in Optional Protocol I Geneva Conventions 1977 which affords protection for the wounded and sick and shipwrecked of any party to a conflict and states that they shall be treated humanely and “shall receive, to the fullest extent practicable and with the least possible delay, the medical care and attention required by their condition”.\textsuperscript{65}

\textsuperscript{60} Art 54 Optional Protocol I Geneva Conventions 1977.
\textsuperscript{61} Mottershaw, E [Supra note 18] Also Verified by the Use of Environmental Modification Techniques. See Lorenz, FM The Protection of Water facilities under international law. http://unesdoc.unesco.org/images/0013/001324/132464e.pdf
\textsuperscript{62} “The State Parties to the present Covenant recognize the right of everyone to the enjoyment of the highest attainable standard of physical and mental health.” Article 12 ICESCR.
\textsuperscript{63} CESCR Gen Comment 14 UN Doc E/C.12/2000/4 para. 30
\textsuperscript{64} Ibid at Para 10.
\textsuperscript{65} Art 10 Protocol I Additional to 1949 Geneva Conventions 1977 http://www.icrc.org/ihl.nsf/f7c4d08d9b287a42141256739003e636b/f6c8b9fee14a77f6c12564e0052b079
Once again the right to health may also be protected in an implicit fashion by IHL, if we were to once again consider Art 54 of Protocol I and the ban on the use of poison that may have serious health effects, or the prohibition of famine as a means of warfare.66 Perrin argues that in relation to the obligation to protect the right to health there will always be a dichotomy between IHL and Human Rights Law and in order to protect the right in times of armed conflict then a balance will need to be struck in order to ensure that Humanitarian considerations are weighed against military necessity and that Human Rights Law can operate to fill in any potential gaps that may arise.67

HOUSING

Article 11(1)68 of ICESCR places an obligation upon state parties to respect the rights of individuals to an adequate standard of living, a right that is inclusive of a right to housing. The committee on Economic, Social and Cultural Rights have also expanded on this right to an adequate standard of living in General Comment 7 where they recognized that practices of forced evictions can occur in situations connected with “forced population transfers, internal displacement [and] forced relocations in the context of armed conflict”.69

The protection of housing in times of armed conflict also receives attention but once again the obligation to protect is much more implicit. The principle of distinction or ‘basic rule’ sets out clearly that “In order to ensure respect for and protection of the civilian population, parties shall... distinguish... between civilian objects and military objects.”70 The prohibition of indiscriminate attack, that is an attack that would strike military objectives, or civilian objects without distinction,71 can also be interpreted in a manner that is consistent with an obligation to protect housing and an adequate standard of living.

67 ibid.
68 “The States Parties to the present Covenant recognize the right of everyone to an adequate standard of living for himself and his family, including adequate food, clothing and housing”
69 CESC General Comment 7 UN Doc E/C.20/05/97. It is perhaps important to note that here the Committee also recognised the protections afforded to civilians under IHL. See para 12.
70 Article 48 Protocol I Additional to the Geneva Conventions 1949.
71 Ibid Art 51 (4).
Along with the evidence provided on the operation of ICESCR in times of armed conflict, its silence on derogation and functionality in territories subject to the control of contracting parties, I submit that the right to water, healthcare and an adequate standard of living all provide evidence that Economic, Social and Cultural Rights are engaged in times of armed conflict and that both International Human Rights Law and International Humanitarian Law can prove to be the means by which such rights are engaged.

**THE GOLDSTONE REPORT**\(^{72}\)

Perhaps one of the greatest reflections on the modernity of economic, social and cultural rights in times of armed conflict comes in the form of the UN Fact Finding Mission to Gaza (Goldstone Report). What is perhaps most interesting about the report, is that it simultaneously applies standards of IHL and Human Rights Law. For instance, when the report was determining upon the legality of the killing of the Hamada Brothers, two prominent businessmen, it applied International Humanitarian law pursuant to Article 54 (1) and (2) of Additional Protocol 1.\(^{73}\) However when determining upon the restrictions that Israel had imposed in Gaza regarding the right to food, the Report used standards of International Human Rights Law, citing the ICCPR as well as Article 12 ICESCR and CEDAW.\(^{74}\) Sometimes the Report used standards of IHL and IHRL in relation to the same events, stating that destruction of housing was not only in violation of the Humanitarian principles contained in Article 147 of the Fourth Geneva Convention regarding a prohibition of destruction on property not justified by military necessity and carried out wantonly, but it was also a violation of Article 11 of ICESCR regarding the right to an adequate standard of living.

There can be a strong argument, as has been made by Pinzauti and Draper, that IHL and IHRL should remain totally separate areas of law. The rationale behind this does not seem too absurd when we see how some confusion may be drawn

---


\(^{73}\) ibid para. 922-934.

\(^{74}\) ibid para 938.
as to exactly what standard applies in times of armed conflict. However, the Goldstone report does provide us with a resounding modern testimony and evidence that both IHL and IHRL are complementary of one another, both contain some of the same provisions and obligations but simply use different language with which to apply them.

CONCLUSION
The International laws of armed conflict and Human Rights have developed in a completely separate manner. At the heart of the two areas of law lies a completely contradictory linguistic framework, one suggestive of peace the other of war. Even at the heart of this contradiction, however, lay deep similarities: both areas of law strive to humanity. I have demonstrated that International Human Rights Law is as equally applicable in times of armed conflict as International Humanitarian Law and that this is inclusive of Economic Social and Cultural Rights. The most recent example of the Goldstone report is proof of this, a report that outlined that even in times of protracted armed conflict, occupation and human rights abuses, the fundamental principles of Economic, Social and Cultural Rights may not be set aside.

MICHEÁL T MURPHY

BIBLIOGRAPHY

BOOKS


JOURNAL ARTICLES


Dennis MJ and Stewart DP *Justiciability of Economic Social and Cultural rights: should there be an international complaints Mechanism to Adjudicate the Rights of Food, Water, Housing and Health.* American Journal of International Law 2004 98 (3) 462-515


Kunnemann, R *A Coherent Approach to Human rights* Human Rights Quarterly 1995 17 (2) 323


Meron, T *The humanization of Humanitarian Law* American Journal of International Law 2000 (4)


Draper, G *Humanitarian Law and Human Rights* Acta Juridica (193) 1979
Lubell, N Challenges in Applying Human Rights Law to Armed Conflict
International review of the Red Cross 2006 (87) 860

Muller, A Limitations to and Derogations from Economic Social and Cultural Rights

Meron, T Extraterritoriality of Human Rights treaties American Journal of International Law 89 (1) 1995

Dennis MJ Application of Human Rights Treaties Extraterritorially in Times of Armed conflict and Military Occupation American Journal of International Law 99 (1) 2005

Cahill, A the Human Right to water – a right of unique status: The Legal status and Normative Content of the right to water International Journal of Human Rights 2005 9 (3)

WEBSITES

http://law.hku.hk/lawgovtsociety/Bangkok%20Declaration.htm
http://www2.ohchr.org/english/law/protocol2.htm
http://www.unhchr.ch/tbs/doc.nsf/0/a5458d1d1bbd713fc1256cc400389e94/$FILE/G0340229.pdf
http://www.unhchr.ch/tbs/doc.nsf/0/94bdaaf5b43a424c12563ed0052b664?opendocument
http://unesdoc.unesco.org/images/0013/001324/132464e.pdf
http://www.icrc.org/ihl.nsf/7c4d08d9b287a42141256739003e636b/f6c8b9fee14a77fdcd125641e0052b079
Defending a General charged with war crimes: 
a summary of the case of Vinko Pandurević at the ICTY

HELENA TOŠIĆ

Background

In the early 1990s, a civil war broke out on the territory of the former Yugoslavia. The worst of the conflict took place in the region of Bosnia and Herzegovina, fuelled by the diverse ethnic composition of the population, comprised of Orthodox Bosnian Serbs, Bosnian Muslims and Catholic Bosnian Croats. In the Eastern part of Bosnia, where the conflict was particularly fierce, the Republic of Srpska was formed on 12 May 1992. Radovan Karadžić, the elected representative of the Bosnian Serbs, became President of the Republic of Srpska and General Ratko Mladić became the Commander of the Army of Republic of Srpska (hereinafter: VRS).¹

Due to the dire humanitarian situation on the ground, the UN Security Council, in April 1993, declared Srebrenica, a town in Eastern Bosnia, a ‘safe area’. The VRS and the Bosnian Muslim Army consequently signed a cease-fire and a demilitarization agreement, which was to be implemented by the beginning of May the same year.²

However, both sides violated the ‘safe area’ agreement. The Bosnian Muslim Army violated the no-fly zone, stationed commands of its units in the town of Srebrenica, and continued to arm itself and conduct military operations within the enclave.³ These actions, particularly the failure to comply with the provisions for the disarmament of the enclave, led the VRS to declare the ceasefire agreement ‘null and void’ and to set itself a goal of demilitarizing the enclave and reducing it to the area defined in the agreement.⁴

¹ Trial Chamber Judgment Prosecutor v Popović et al., 10 June 2010, paras. 88-90.
² Ibid., paras. 92-97.
³ Ibid., para. 98; See also Prosecution Adjudicated Facts Decision Prosecutor v Popović et al., Annex, Facts 46 and 50. A The United Nations Military Observers (UNMO) were also deployed in the Srebrenica enclave, tasked with monitoring violations of the ceasefire agreement. See Trial Chamber Judgment Prosecutor v Popović et al., 10 June 2010, para. 191 et seq.
⁴ Ibid., para. 99.
In early July 1995, the VRS launched an attack against the Bosnian Muslim Army deployed in the town of Srebrenica, and on 11 July 1995, the VRS took control of the town. Following the fall of Srebrenica, thousands of Bosnian Muslims attempted to flee the area. The population gathered in the small town of Potočari. There, members of the VRS separated Bosnian Muslim men from their families, and whilst women, children and elderly people were transferred to the territory controlled by the Bosnian Muslim Army, many of the able-bodied Muslim men were eventually detained, transported to remote locations near towns of Bratunac and Zvornik, and summarily executed.

**Popović et al. case at the ICTY**

The International Criminal Tribunal for the Former Yugoslavia (ICTY) based in The Hague has held trials for over 21 high-ranking military personnel from the VRS for the events that took place in Srebrenica, with the largest trial, in the case of *Prosecutor v Popović et al.*, having included seven co-defendants. The indictment included a number of counts and nearly all of the accused were charged with individual and command responsibility for genocide, conspiracy to commit genocide, extermination, persecution, forcible transfer, deportation and murder as a war crime and a crime against humanity. These crimes were said to have been committed in furtherance of two common plans — a joint criminal enterprise to murder able-bodied Muslim men from Srebrenica and a joint criminal enterprise forcibly to remove the Bosnian Muslim population from Srebrenica — which, together, were said to amount to a crime of genocide.

The trial in the *Popović et al.* case took three years to complete and is considered one of the largest and most complex international criminal trials to have taken

---

7 *Indictment Prosecutor v Popović et al.*, IT-05-88-T, 4 August 2006, paras. 26-95. These crimes are punishable under the Articles 4(3)(a) (genocide), 4(3)(b) (conspiracy to commit genocide), 5(b) (extermination), 5(h) (persecutions), 5(i) (forcible transfer), 5(d) (deportation) and 3 (murder as war crime and crime against humanity) of the Statute of the Tribunal. For additional requirements that must be satisfied before a crime falls within the ICTY’s jurisdiction under Article 3 of the Statute see Trial Chamber Judgment *Prosecutor v Popović et al.*, 10 June 2010, paras. 739-743.
place since the Second World War. Testimonies from over 300 witnesses were presented and almost 90,000 pages of exhibits reproduced,\(^9\) including complex military manuals, war logbooks, intercepted communications, satellite and aerial images, war maps, video footages and reports of demographic, anthropological, pathological and other experts.\(^{10}\)

**General Vinko Pandurević and his role in the events**

General Vinko Pandurević, one of the co-defendants in the *Popović et al.* case, was a brigade commander in the area of Zvornik, a town close to Srebrenica, during the war. At that time he held the rank of Lieutenant Colonel and was, at the age of 33, in command of over 5,000 members of the Zvornik Brigade.\(^{11}\) The majority of the members of his Brigade were so-called ‘weekend warriors’, i.e. local farmers, factory workers and others that held regular jobs but took part in defending Serbian positions and forming part of the Brigade for two weeks every month. Commanding such a brigade was an extremely cumbersome task — Pandurević often faced resistance and many of his attempts to organise the Brigade effectively ended in vain.\(^{12}\)

In July 1995, Pandurević was appointed as one of the officers to take part in the attack on Srebrenica aimed at the demilitarization of the enclave,\(^{13}\) and was designated to command a special military unit, Tactical Group 1 (TG-1).\(^{14}\) While Pandurević was absent from Zvornik and involved in pursuing what were, to a degree, lawful military objectives of fighting the enemy as the commander of TG-1, his deputy commander took over the combat operations of the Zvornik Brigade.\(^{15}\) It was during Pandurević’s absence that the high-ranking VRS officers

---

\(^{9}\) Trial Chamber Judgment *Prosecutor v Popović et al.*, 10 June 2010, para. 5.

\(^{10}\) Communications intercepted and transcribed by the Bosnian Muslim Army and aerial images were particularly challenged by several Defence teams. See Trial Chamber Judgment *Prosecutor v Popović et al.*, 10 June 2010, paras. 64-66, 72-75. For further reading on the admissibility of evidence at the ICTY see Khan, Buisman, Gosnell: *Principles of Evidence in International Criminal Justice*, Oxford University Press, 2010.

\(^{11}\) Trial Chamber Judgment *Prosecutor v Popović et al.*, 10 June 2010, paras. 143, 1839-1841.


\(^{13}\) Trial Chamber Judgment *Prosecutor v Popović et al.*, 10 June 2010, para. 1843.

\(^{14}\) Trial Chamber Judgment *Prosecutor v Popović et al.*, 10 June 2010, para. 1843.

\(^{15}\) *Ibid.*, para. 1842; Sentencing Judgment *Prosecutor v Obrenović*, IT-02-60/2-S, 10 December 2003; See also VRS Rules on Brigade Commanders Authority of the Regiment, Article 17: ‘In the absence of
brought large numbers of Bosnian Muslim prisoners to the area of Zvornik and executed them. These officers employed and used men on the ground, including a few members of the Zvornik Brigade to assist them in guarding, executing and burying prisoners.\footnote{Ibid., paras. 1861, 2016-2020.}

**Pandurević Defence case**

The preparation of the Defence case for Pandurević was a challenging and lengthy process, stretching over a period of almost four years. Pandurević was the only accused to have testified in his own case in the Popović *et al.* trial. He presented extensive evidence and underwent vigorous cross-examinations by other Defence teams as well as by the Prosecution, lasting over a month.

One of the most important achievements of the Pandurević Defence team was to deconstruct and discredit the Prosecution’s theory about the so-called ‘area of responsibility’. The Prosecution argued, in effect, that as the commander of the Zvornik Brigade, Pandurević was responsible for the territory of Zvornik and was therefore ultimately accountable for all the events that took place on that territory.\footnote{Prosecution Final Trial Brief *Prosecutor v Popović et al.*, Public Redacted Version, 14 July 2010, paras. 484, 845, 945, 1072-1077.} However, the Defence was able to prove that, under customary international law, a military commander can only be held responsible for the acts of his subordinates, and never for the territory — as a general rule, it is the State, not the individual military commander, that is responsible for a certain territory.\footnote{The only exception are occupation commanders or military governors endowed with executive powers over a territory occupied by their forces and a general duty to ensure the well-being of the civilian population within that territory. See for example ‘The Hostage Case’, *United States v. Wilhelm List et al.*, Law Reports of Trials of War Criminals, UN War Crimes Commission, Vol. VIII, pp. 69-70.}

Another fundamental objective of the Defence was to show that the principle of ‘singleness of command’ meant that a brigade cannot be under the command of two commanders at the same time, as well as that an individual can only command one military unit at a time. When the murder operation moved to the
Zvornik area, Pandurević was not only absent and out of contact with the Zvornik Brigade, but was in command of the TG-1, a different military unit that participated in what were, at least partly, lawful operations to demilitarize the Srebrenica enclave. Accordingly, it must have been the Pandurević’s deputy commander who was in charge of the Zvornik Brigade at the critical time and thus responsible for those brigade soldiers that were sent by the superior VRS officers to participate in the murder operation.¹⁹

Further, the Defence argued that Pandurević had no necessary ‘effective control’ over all Zvornik Brigade members at the critical time. Although the ‘superior order’ defence is not accepted under international customary law²⁰, this must be distinguished from a situation that took place in Zvornik in July 1995, where a senior high-ranking military officer surpasses the regular military chain of command and deploys a brigade soldier on the ground without first contacting his brigade commander. The VRS Main Staff officers came to the area and ordered the ‘weekend warriors’ to take part in the detention and burying of prisoners at the moment when their de jure commander, Pandurević, was absent.²¹ The Defence argued that it would be wrong to hold Pandurević responsible for such dysfunctioning of the regular command structure.

In addition, the Defence also sought to establish the absence of genocidal intent on the part of Pandurević. As he returned to the Zvornik area and became aware of the terrible events which took place during his absence, Pandurević contacted the commander of the Bosnian Muslim Army and opened a corridor allowing the remainder of the Bosnian Muslim soldiers and civilians to pass through and walk to the free territory: it is estimated that up to 6,000 Bosnian Muslims safely passed through the corridor and escaped the murder operation.²² Notably,

---

²⁰ For an overview of the development of this principle see Antonio Cassese: International Criminal Law, Oxford University Press, 2008, Chapter 13.
²² Ibid., paras. 551-561.
Pandurević took this action in contravention of the orders from his superiors, including General Mladić.

Judgment and dissenting opinion of Judge Kwon
The eagerly awaited judgment in Popović et al. was delivered on 10 June 2010. The Trial Chamber ruled that the events in and around Srebrenica in July 1995 amounted to genocide and passed the most serious sentences to the accused, including life imprisonment.23

In respect of Pandurević, the Trial Chamber held that during his absence from Zvornik, Pandurević had minimal contact with the Brigade and that there was 'no evidence that he was aware of the events transpiring in the Zvornik area, including the murder operation, let alone that he contributed to them'.24 Although some of the victims were executed and buried after Pandurević had returned to the Zvornik Brigade, the Prosecution nonetheless failed to provide any evidence that Pandurević himself participated in, or ordered, authorised or otherwise approved the participation of his subordinates in that murder operation.25

The Trial Chamber further held that it was not established beyond reasonable doubt that Pandurević participated in a joint criminal enterprise to murder or a joint criminal enterprise forcibly to transfer the Bosnian Muslim population. He was, accordingly, acquitted of the charges of genocide and exterminations.26

Furthermore, the Trial Chamber considered that while Pandurević’s role in the attack on Srebrenica assisted in the forcible removal of the Bosnian Muslim population from the enclave, this was in part a justifiable military operation to demilitarize the enclave. The limited nature of Pandurević’s involvement in the forcible transfer diminished the gravity of his criminal conduct and this was

23 Ibid., para. 2227 and Chapter IX. Disposition.
24 Ibid., para. 1969.
taken into account in determining his sentence for the crime of aiding and abetting forcible transfer.\textsuperscript{27}

Moreover, the Trial Chamber found that after Pandurević’s return to Zvornik, there was a single incident when ten Bosnian Muslim soldiers had been taken by a superior VRS officer to be executed. Although the Chamber found no evidence of any positive acts on Pandurević’s part that may have aided the murder of these ten men, it nonetheless concluded that Pandurević had failed to discharge his legal duty to protect the wounded prisoners. Accordingly, the majority (Judge Kwon dissenting) found Pandurević guilty of aiding and abetting the crime of murder of ten Muslim prisoners by omission.\textsuperscript{28}

The majority (Judge Kwon dissenting) also held that Pandurević was in command of the Zvornik Brigade even during his absence from the area and while in command of the TG-1. Given that the test for the superior-subordinate relationship rests on the \textit{ability} to effectively control (as opposed to the \textit{exercise} of that control), there is no exclusivity in the determination of effective control, as opposed to the principle of the ‘singleness of command’. Additionally, the majority found Pandurević guilty as a commander for failing to prevent further crimes upon his return.\textsuperscript{29}

In determining the sentence for Pandurević, the Trial Chamber stressed the importance of the uncommon and extraordinary set of circumstances that he had faced at the critical time. In particular, Judge Kwon was struck by the strong and defiant language used by Pandurević in his combat reports addressed to the superior command at that time. In Judge Kwon’s view, these reports were ‘\textit{the sole instance of a subordinate so openly challenging Mladić in relation to murder operation.’}\textsuperscript{30} Pandurević’s bravery in these reports and his actions ‘\textit{evidence his}

\textsuperscript{27} \textit{Ibid.}, paras. 2211-2212.
\textsuperscript{28} \textit{Ibid.}, paras. 1981-1991; See also Appeal Judgment \textit{Prosecutor v Mrksić and Šljivančanin}, paras. 73, 151: The breach of the duty to protect prisoners of war as imposed by the laws and customs of war may give rise to individual criminal responsibility.
\textsuperscript{29} \textit{Ibid.}, paras. 2012-2066.
\textsuperscript{30} Trial Chamber Judgment \textit{Prosecutor v Popović et al.}, 10 June 2010, Dissenting Opinion of Judge Kwon, para. 79; See also the similar wording used by the majority in para. 2221.
character, namely his strength and integrity, as an individual and as a commander. In his dissenting opinion, Judge Kwon stated that the sentence imposed by the majority does not adequately account for the nature of Pandurević’s involvement and the mitigating circumstances and should therefore be further substantially reduced.

The majority sentenced Pandurević to 13 years imprisonment and concluded that his decision, taken against the orders of his superiors, to open the corridor to enable the safe passage of thousands of Bosnian Muslim men was ‘striking’ and saved the lives of thousands of men. His action in this regard ‘stands out as an instance of courage and humanity in a period typified by human weakness, cruelty and depravity’.

Conclusion
Pandurević’s case is currently on appeal and the Prosecution and the Defence are seeking the Appeal Chamber to overturn some of the legal and factual findings in his case.

Whilst the importance of prosecuting and punishing the perpetrators of such brutal crimes as those committed in Srebrenica cannot be overstated, it is nonetheless the professional duty of every lawyer, including prosecuting counsel, to recognize that not all of those accused before the ICTY deserve the highest possible sentence. If the aim of the Tribunal is to bring justice to the victims, it needs to distinguish between those that were giving orders on the ground and those whose acts were a clear and compelling instance of assistance to potential victims. It is in that sense that the Defence hopes that the Appeal Chamber will hold that those, whose lives were in Pandurević’s hands, have survived.

HELENA TOŠIĆ

31 Trial Chamber Judgment Prosecutor v Popović et al., 10 June 2010, Dissenting Opinion of Judge Kwon, paras. 78-80.
32 Ibid., para 81.
33 Trial Chamber Judgment Prosecutor v Popović et al., 10 June 2010, para. 2219 and Chapter IX. Disposition.
Indus Water Treaty in the ICJ: An Alternative Approach to Resolve Indo-Pak Water Disputes

AQEEL MALIK

Introduction

This paper will discuss and analyse how the water disputes between India and Pakistan would have been decided based on the assumption that both states had submitted this contentious matter to the jurisdiction of the International Court of Justice (ICJ), referring their disputes to be decided by this Court in accordance with the norms of international law. This approach is unique in the sense that some scholars and experts have spoken about how the disputes cannot be referred to the ICJ in light of the dispute resolution provisions of the Indus Water Treaty (IWT), but none of them have discussed the slightest possibility as to how the decision would turn out if it had come before the ICJ for adjudication. This paper, therefore, will outline the disputes since the IWT came into force, in addition to focusing on how the ICJ would have decided these disputes.

The Indus Water Treaty

‘The Indus Waters Treaty (IWT) is one of the most shining examples of dispute resolution because it contributes not only to the solution of an international dispute but also to the development of a scheme’. The treaty is considered to be a complex instrument with a fundamental approach to divide the water resources of the Indus Basin equally between India and Pakistan and to increase the quantity of water available to both the states. This is the only international water treaty that is signed by a third party. Being one of the most important rivers in the world, the Indus River is situated in Northwest India and Pakistan. The Indus system is made of up the main Indus River and its major tributaries. The Indus Basin begins in the Himalayan Mountains in the State of Jammu and Kashmir flowing from the hills through the arid states of Punjab and Sindh and

---

1 The Indus Water Treaty of 1960.
2 For example, Anna Spain.
6 R.K. Arora, supra at 9.
emptying into the Arabian Sea. When the partition took place, most of the water-rich areas went to India. As a result, Pakistan became the water-short lower riparian.\textsuperscript{7} IWT was signed in Karachi, Pakistan in 1960 by the Prime Minister of India, Jawaharlal Nehru, and the President of Pakistan, Mohammad Ayub Khan.\textsuperscript{8} It was signed with the help of the ‘good office’ of the World Bank after mediating nearly a decade of negotiations.\textsuperscript{9} However, just before the World Bank intervened, both India and Pakistan had signed the Inter-Dominion Accord of May 1948 designed to meet their immediate requirements. In pursuit of a permanent solution to this problem, Pakistan had proposed that the dispute be referred to the ICJ for adjudication but India disagreed and put forward the idea of an \textit{ad hoc} bilateral tribunal.\textsuperscript{10} Moreover, there is a contention that the treaty was somewhat forced upon both the states.

The main purpose behind drafting this treaty was to clearly outline the rights and responsibilities of each state’s water use against each other.\textsuperscript{11} The treaty contains twelve articles and eight annexes that also include the appendices dealing with issues of allocation of water and its flow.\textsuperscript{12} Furthermore, the treaty divided the Indus River system into Eastern and Western Rivers.\textsuperscript{13} India was allocated the Eastern Rivers that include Ravi, Beas, and Sutlej, whereas, Pakistan was given the Western Rivers that include Indus, Chenab, and Jhelum. The Eastern Rivers allocated to India were subject to a 10-year transitory period in which the supply of a certain quantum of water to Pakistan had to be guaranteed whilst Pakistan carried out the required construction works on the Western Rivers in replacement of its Eastern Rivers sources.\textsuperscript{14} Contrary to this, Pakistan’s utilisation of Western Rivers was subject to India’s right to make use of some of the water on these rivers before they entered Pakistani territory for irrigation purposes, generating hydroelectric power and other assigned

\textsuperscript{7}Salman and Boisson de Chazournes, \textit{International Watercourses Enhancing Cooperation and Managing Conflict}, Ch. 9, Pg. 155, (1998).
\textsuperscript{8}IWT 1960, Preamble.
\textsuperscript{9}Id.
\textsuperscript{10}Salman and Boisson de Chazournes, \textit{supra} at 159.
\textsuperscript{11}R.K. Arora, \textit{supra} at 10.
\textsuperscript{12}IWT 1960.
\textsuperscript{13}Id. Article I.
\textsuperscript{14}R.K. Arora, \textit{supra} at 10-11.
functions.\textsuperscript{15} In essence, India has to abstain from interfering with the water of Western Rivers, except that it is allowed to use the water from these rivers for domestic, non-consumptive, power generation, and limited agricultural purposes.\textsuperscript{16} Likewise, Pakistan under the treaty is permitted to make use of the water from the Eastern Rivers for non-consumptive, domestic, and certain agricultural purposes. However, it has to refrain from impeding the water of Eastern Rivers until they have crossed into Pakistani territory.\textsuperscript{17}

Interestingly, the treaty permits both states to independently develop hydropower and this depicts the treaty to be ‘an example of a water trade-off, rather than water sharing’.\textsuperscript{18} Some scholars believe that ‘the fact that the treaty does not specify the quantity of water allocated means that it effectively provides for a territorial type of sharing’, though, not true it is a division that concerns only the use of the water.\textsuperscript{19} Besides, IWT also provides for assistance and collaboration between both states as it recognises that there is a mutual interest in the most favourable growth of the rivers.\textsuperscript{20}

**Dispute Resolution Under The Treaty**

The treaty provisions provide a detailed and complex mechanism for resolving a dispute that may arise between both the parties.\textsuperscript{21} In the first instance, the Permanent Indus Commission (PIC)\textsuperscript{22} examines all of the disputes and it tries to resolve them.\textsuperscript{23} The PIC is formed of two Commissioners, one from each state, and its prime aim is to implement the treaty whilst establishing and upholding mutual provisions.\textsuperscript{24} There is a requirement for the Commission to meet at least once a year, alternating between India and Pakistan, and when either of the Commissioners requests so.\textsuperscript{25} Either Commissioner may vouch that the issue

---

\textsuperscript{15} Id.  
\textsuperscript{16} Id.  
\textsuperscript{17} Id.  
\textsuperscript{18} R.K. Arora, supra at 13.  
\textsuperscript{19} R.K. Arora, supra at 23-24.  
\textsuperscript{20} Id. at 13.  
\textsuperscript{21} IWT 1960, Article IX (Settlement of Differences and Disputes).  
\textsuperscript{22} IWT 1960, Article VIII.  
\textsuperscript{23} Id. at Para 1.  
\textsuperscript{24} R.K. Arora, supra at 14.  
\textsuperscript{25} Id. at 15.
comes within one of the 23 selected areas for it to be treated as a ‘difference’ if the Commission is unable to concur. The ‘difference’ will then have to be dealt with by a neutral expert who ought to be a highly qualified engineer and appointed by the World Bank.\(^{26}\) However, if the ‘difference’ does not come under one of the 23 selected areas or the neutral expert renders that the ‘difference’ be treated as a ‘dispute’ then should the governments of both states desire to mediate, they may do so.\(^{27}\)

Furthermore, the last resort to dispute resolution provided for under the treaty is arbitration. The Court of Arbitration should consist of seven members and they are to be appointed in consultation with the World Bank. The dispute may be referred to arbitration by agreement between both the states or at the request of either state if they believe that the dispute is unlikely to be resolved by other means of resolution or an undue delay may be caused in conducting negotiations.\(^{28}\)

**Breaches of IWT Provisions**

Pakistan contends that India has breached a number of provisions stated in the treaty and is still continuing to violate them. The alleged breaches are with regards to different hydroelectric projects that India has and is still constructing in its territory. These projects and alleged breaches will be discussed below:

1. **Salal Hydroelectric Project:** This project is located on River Chenab and is a run-of-river project.\(^{29}\) Pakistan was first informed of this project in 1974 and it objected to the design of this project that included ‘six low-level outlets and overall height of spillway gates of 40 ft’ in clear violation of the treaty.\(^{30}\) However, the issue was resolved bilaterally between both

---

\(^{26}\) **IWT 1960, Article IX.**
\(^{27}\) **R.K. Arora, supra** at 16.
\(^{28}\) **Id.**
\(^{29}\) **Azhar Ahmad, Indus Water Treaty A Dispassionate Analysis,** Institute of Policy Studies, at 4, (2011).
\(^{30}\) **Id.**
the countries in 1978 by an agreement to reduce the height of the spillway gates from 40 to 30 ft that proved to be successful.31

2. **Baglihar Hydroelectric Power Project**: It is located upstream of Salal project on River Chenab and is also a run-of-river project.32 This project was approved in 1996 and construction started in 1999.33 The features of this project include a high concrete gravity dam with a gross storage and power generation capacity. Permission was granted to the Pakistani team in 2003 for site inspection.34 Objections related to the design of undersluice type gated spillway were raised by Pakistan in violation of the treaty.35 This could potentially cause shortage of inflows during lean months if the dam is mal-operated by India as in light of the treaty; Pakistan contends that the design ought to be based on ‘un-gated spillways’.36

The objection raised by Pakistan could not be resolved by the PIC and the issue was referred for expert determination.37 Furthermore, Mr. Raymond Lafitte who was appointed as a neutral expert by the World Bank in 2005 gave his Final Determination in 2007.38 The decision by the neutral expert upheld some of Pakistan’s contention asking India to reduce free board from 4.5m to 3m but did not ask to change the spillway design.39 As a result of this, India obstructed River Chenab’s flow whilst filing the dam in 10 days instead of spreading it out over the span of 60 days thereby violating the treaty provisions.40 This resulted in loss of water flow received by Pakistan that greatly affected its agriculture sector even after numerous requests to India to provide Pakistan with the filling schedule.41 The issue remains unresolved as Pakistan has sought compensation from India for the lost waters, however,

---

32 Azhar Ahmad, *supra* at 4.
33 Id.
34 Id.
35 IWT 1960, Para 8, Annexure D.
36 Azhar Ahmad, *supra* at 5.
37 Id.
38 Id.
39 Id. at 5.
40 Id.
41 Id.
India is inflexible with regards to this matter and Pakistan has demanded monitoring of River Chenab’s water flow by setting up telemetry device.\[^{42}\]

3. **Wullar Barrage:** It is located on River Jhelum in Indian Occupied Kashmir and is also called ‘Tulbul Navigation Project’. The project was first initiated in 1984 without informing Pakistan as it is required under the treaty and Pakistan came to know about this project in 1986.\[^{43}\] India would be able to gain regulation control of the Wullar Lake by creating supplementary storage volume that would aide in amassing water for up to 6 months during the wet season.\[^{44}\] The importance of this project is that it could potentially destroy the intact ‘triple canal project’ of Pakistan.\[^{45}\]

Construction works were suspended by India in 1987 after Pakistan strongly protested against the features of this project as according to the treaty, India is allowed to construct such projects on Main Chenab and Jhelum Rivers with a total storage capacity of not more than 10,000 acre ft.\[^{46}\] Whereas, the current storage capacity of the barrage is 30,000-acre ft that being at least three times in excess of the allowance.\[^{47}\]

Furthermore, Pakistan argues that this project is a major breach of **Article I (ii)** of the IWT and the construction of such projects by India is only allowed after Pakistan has examined and approved the design of the project.\[^{48}\] Accordingly, both the countries are currently resolving the issue and works are discontinued.\[^{49}\]

\[^{42}\] Id.
\[^{44}\] Azhar Ahmad, *supra* at 5.
\[^{45}\] Id.
\[^{46}\] IWT 1960, Annexure E, sub-para 8 (h).
\[^{47}\] Azhar Ahmad, *supra* at 5.
\[^{48}\] Id.
\[^{49}\] Id.
4. **Kishanganga Hydroelectric Project**: this is another hydroelectric project being constructed by India on the Neelum River.\(^{50}\) India provided information regarding this project to Pakistan in June 1994 when the construction of a tunnel started on Neelum River to divert the waters into the Wullar Lake.\(^{51}\) The project is designed to divert the flow of this dam into Wullar Lake through a 21.6 KM tunnel after generating 330 MW of power. Under the treaty, India is allowed legal construction of storage works for up to 0.75 million acres feet (MAF) on the Jhelum River tributaries.\(^{52}\)

Pakistan objected to the construction of this project and contended on two points.\(^{53}\) The first objection is with regards to the design criterion, as it does not conform to the treaty provisions.\(^{54}\) The second objection is the diversion of flow of one tributary into another resulting into a flagrant breach of the treaty.\(^{55}\) The impact of this project on Pakistan would be depriving it of its own hydroelectric project called ‘Neelum-Jhelum’ that would result in some crucial emancipation by reducing the production capacity from 11 to 16 percent.\(^{56}\) This would also interfere with the agriculture uses in the Neelum Valley apart from having a direct effect on the socio-economic life and ecological aspects in the area. Furthermore, the matter is unresolved due to the failure of both Pakistan and India to resolve this issue bilaterally after having five meetings on this particular issue since November 2004. Pakistan took this issue to the Permanent Court of Arbitration (PCA) to try and resolve it there as provided under the treaty.\(^{57}\) The court handed down interim measures preventing India from continuing permanent construction works.\(^{58}\)


\(^{51}\) Id.

\(^{52}\) IWT 1960, Annexure E, Para 7.

\(^{53}\) Azhar Ahmad, *supra* at 6.

\(^{54}\) IWT 1960, Annexure E, Para 11.

\(^{55}\) IWT 1960, Annexure E, Para 10 “Diversion of flow of one tributary into another of River Jhelum is not provided for in this annexure”.

\(^{56}\) Azhar Ahmad, *supra* at 6.

\(^{57}\) IWT 1960, Article IX (2).

5. **Dul Hasti Hydroelectric Plant:** This is a 390 MW under construction project on River Chenab. The construction began in 1999 and the first phase of this project was completed in 2007 after being initiated in 1983. This hydroelectric plant is a ‘high concrete gravity dam’ upstream of Baglihar hydroelectric project on River Chenab. Though the adverse effects of this particular project on Pakistan are not severe because the water can only be blocked for 1 to 2 days, nevertheless it is crucial to dissuade India from providing ‘under-sluices type gated spillways in the body of the dam’.

6. **Bursar Dam:** This project is currently under construction and is being built on both River Jhelum and Chenab flowing through Indian Occupied Kashmir. It is also anticipated to be a major project that would preserve 2.2 MAF and generate 1020 MW of electricity. This dam is considered a storage facility that is anticipated to be completed over 6-7 years and this would regulate the water flow to downstream projects. Moreover, the current storage capacity of 2.2 MAF is considered beyond the limits permitted under the treaty. Not only does this particular project violate the treaty, but it also violates international environmental conventions. This project does not only cause water scarcity in Pakistan but also aides to be a contributing factor for melting of Himalayan glaciers.

7. **Nimoo Buzgo Hydropower Project:** This is a run-of-river hydroelectric project on River Indus with an installed capacity of 45 MW. Pakistan learnt about this project in 2002 and requested India for further

---

60 Azhar Ahmad, *supra* at 6.
61 Id.
62 Id.
63 Id.
64 Id.
65 Id.
66 Id.
68 Azhar Ahmad, *supra* at 6.
information relating to this project but until December 2006, India failed to respond despite numerous reminders.\textsuperscript{70} After Pakistan inspected the site, it learnt that the project is now 80 percent complete and is due to be inaugurated in August 2012.\textsuperscript{71} Furthermore, Pakistan has already decided to take this issue to the Permanent Court of Arbitration (PIC)\textsuperscript{72} and get it resolved there as India has flagrantly breached the IWT by not informing, seeking approval and/or getting the project’s design scrutinized prior to starting construction works.

**Comparable Environmental Law Issues Decided by the ICJ**

So far, ICJ has only decided two environmental law cases relating to the use of international water recourse that are relevant to the issue between Pakistan and India in its entire history.\textsuperscript{73} To deal with international environmental law issues, ICJ established a special chamber of the Court for environmental matters in 1993 pursuant to Article 26 (1) of its Statute, however, it was discontinued in 2006 but it did raise the question of special international adjudicatory forum for environmental disputes.\textsuperscript{74} Both these cases will be discussed below:

1. **Pulp Mills on the River Uruguay (Argentina v. Uruguay)** – the summary of this case is that “In 2003, Uruguay authorised the construction of pulp mills on the Uruguay River. In 2006, Argentina initiated proceedings to prevent Uruguay from constructing the mills. Argentina claimed that Uruguay violated treaty provisions that require prior notification and consultation before taking actions that could affect river water quality. The Court rejected Argentina’s request for a preliminary injunction, stating that Uruguay intended to and could still comply with its international obligations. As a result, protesters in Argentina blockade roads to prevent construction. When

\textsuperscript{70} Id.
\textsuperscript{71} Id.
\textsuperscript{74} International Court of Justice, Composition of the Chamber for Environmental Matters, (March 2 2002), http://www.icj-cij.org/presscom/index.php?pr=106&p1=6&p2=1&search=%22%22Composition+of+the+Chamber+for+Environmental+Matters%22%22.
Argentina appealed to the Court a second time, Uruguay also sought relief from Argentina’s protests. The Court rejected both requests because neither state presented risks of prejudice to their rights under the treaty. On 20 April 2010, the Court concluded that while Uruguay breached its international procedural obligations to notify and consult with Argentina before authorizing and commencing construction on the pulp mills, the Court’s declaration of Uruguay’s breach constituted a sufficient remedy for Argentina’s claim.\textsuperscript{75}

2. **Case Concerning the Gabcikovo-Nagymaros Project (Hungary v. Slovakia)** – the summary of this case is “In 1977 Hungary and Czechoslovakia signed a treaty obligating the States to cooperate in the construction of a system of dams and locks along a section of the Danube River that formed the border between the States. Construction commenced in 1978 but progressed slowly due to political and economic transformations in both States. In 1989, Hungary abandoned the project, justifying its decision on claims of changed circumstances and impossibility. In 1993, Czechoslovakia peacefully separated into two nations: Czech Republic and Slovakia. Slovakia assumed its predecessor’s responsibilities under the treaty because the planned hydraulic system fell within its territory along the Danube River. After continued negotiations failed, Slovakia devised "Variant C," an alternative plan to complete the project. Under Variant C, Slovakia dammed the Danube and appropriated between 80 and 90% of the river water. The dispute came before the International Court of Justice in 1994 and was decided in 1997. The Court rejected Hungary’s claims of changed circumstances and impossibility but also concluded that Slovakia, by putting Variant C into operation and unilaterally taking control of a shared resource, had violated international law and the 1977 Treaty. Ultimately, the Court ordered the parties to "re-establish co-operative administration of what remains of the Project."\textsuperscript{76}

\textsuperscript{75}International Court of Justice, at \url{http://www.icj-cij.org/docket/index.php?p1=3&p2=3&k=88&case=135&code=au&p3=5}

\textsuperscript{76}Ibid. at \url{http://www.icj-cij.org/docket/index.php?p1=3&p2=3&k=8d&case=92&code=hs&p3=5}

This convention was adopted by the United Nations General Assembly on May 21, 1997 and is considered to be the first global water law.77 A majority of States numbering up to 103 voted in favour of this Convention and only 3 States voted against it.78 Only 27 States abstained, including Pakistan and India.79 The most significant articles of this convention are Article 5 and 780 and these will be discussed below with reference to the IWT.

Firstly, Article 5 of the Convention states:

“Watercourse States shall in their respective territories utilize an international watercourse in an equitable and reasonable manner. In particular, an international watercourse shall be used and developed by watercourse States with a view to attaining optimal and sustainable utilization thereof and benefits therefrom, taking into account the interests of the watercourse States concerned, consistent with adequate protection of the watercourse. Watercourse States shall participate in the use, development and protection of an international watercourse in an equitable and reasonable manner. Such participation includes both the right to utilize the watercourse and the duty to cooperate in the protection and development thereof, as provided in the present Convention”.81

The principle set forth in this article is the equitable and reasonable utilisation of an international watercourse.82 According to this article, the use must be coherent with sufficient protection of the watercourse from dilapidation including pollution.83 This is evident from the ICJ’s decision in Gabcikovo-

---

77 Salman and Boisson de Chazournes, International Watercourses Enhancing Cooperation and Managing Conflict, Ch. 2, Pg. 17, (1998).
78 Id.
81 Id. at Article 5.
82 Salman and Boisson de Chazournes, supra at 19.
83 Id.
Nagymaros in which the court stressed on the watercourse to be used ‘in an equitable and reasonable manner’. Additional, this article also initiated the concept of ‘equitable protection’, which essentially encompasses the idea of riparian states to take positive steps whilst cooperating with each other with regards to the watercourse.

**Article 7** is the most controversial provision of this convention as it states: “Watercourse States shall, in utilizing an international watercourse in their territories, take all appropriate measures to prevent the causing of significant harm to other watercourse States. Where significant harm nevertheless is caused to another watercourse State, the States whose use causes such harm shall, in the absence of agreement to such use, take all appropriate measures, having due regard for the provisions of articles 5 and 6, in consultation with the affected State, to eliminate or mitigate such harm and, where appropriate, to discuss the question of compensation.”

This Article is thought to be controversial due to its relationship with **Article 5** of the same convention. The controversy is one of precedence between both the articles and which one should prevail over the other and it seems that precedence is given to **Article 5** that encompasses ‘equitable utilisation’ over **Article 7** enlists ‘no harm doctrine’. However, in light of **Article 10** of this Convention that provides the mechanism for resolving conflicts in using an international watercourse states that it must be done so with reference to **Article 5 to 7**. This means that the conflict is not to be resolved solely on the basis of ‘no harm doctrine’ set forth in **Article 7**, instead, it is to be resolved in light of all the Articles in this Convention that include both principles of ‘equitable utilisation’ and ‘no harm doctrine’.

---

84 Id.
85 Id. at 20.
87 Salman and Boisson de Chazournes, supra at 21.
88 Id.
89 Id. at 22.
90 Id.
91 Id.
IWT clearly states in its provisions the need for equitable utilisation and not to cause harm to the riparian states.\textsuperscript{92} However, India in using water from these rivers has been in breach of the treaty provisions on various occasions with regards to the current and previous projects that it initiated on these rivers. In light of the ‘equitable utilisation’ principle as stated in \textbf{Article 5} of the Convention, India has at times not utilised the water equitably, depriving Pakistan from carrying out its agriculture and other important activities at an optimum level, thereby, causing harm to Pakistan by making use of the international watercourse the way it has so far and is still continuing to do so.

Although 103 states voted with assent to the Convention and this demonstrates that there is a general acceptance of the Convention, there is also an argument that recent environmental disputes have not been successfully dealt with, which renders it to be a weak legal instrument for resolving conflicts.\textsuperscript{93} Additionally, this Convention has not been ratified by UN membership as of January 2009.\textsuperscript{94} On the contrary, I am of the view that this Convention should have been signed and ratified by both India and Pakistan instead of abstaining as IWT was signed in 1960 and the Convention was adopted in 1997. The Convention in comparison to the IWT is recent and addresses recent environmental challenges not present at the time the treaty was signed. The Convention can be used, as an additional legal instrument to complement the treaty with regards to recent changes in our environment and water needs. This is also evident from the ICJ’s decision in the \textit{Gabcikovo-Nagymaros} case, as it placed great emphasis on the fact that difficult cases will be resolved by cooperation and compromise and not by inflexible assertion on the rules.\textsuperscript{95}

\section*{How Would The ICJ Adjudicate the Issues under IWT?}

International Court of Justice (ICJ) is the principal judicial organ of the United

\textsuperscript{92} IWT 1960, Article II, III, and IV.
\textsuperscript{93} Wikiwater.net, Negotiations on the Convention, October 17 2009, at \url{http://waterwiki.net/index.php/Convention_on_the_Law_of_the_Non-Navigational_Uses_of_International_Watercourses}
\textsuperscript{94} Id.
\textsuperscript{95} Salman and Boisson de Chazournes, \textit{supra} at 22.
Nations (UN) and all member states of the UN are *ipso facto* parties to the Statute of the Court.\(^{96}\) The Court’s jurisdiction is divided into two distinctive types; namely, contentious and advisory jurisdiction and the IWT issued will fall into contentious jurisdiction as it only applies to disputes between states having agreed to the jurisdiction.\(^{97}\)

Notwithstanding any dispute resolution provisions prescribed in the IWT and proceeding to the ICJ based on the assumption that both India and Pakistan have accepted the Court’s jurisdiction on an *ad hoc* basis under **Article 36 (1)**,\(^{98}\) the ICJ will have jurisdiction to adjudicate on the disputes arising under the treaty. However, some scholars argue that ICJ has been hesitant to proffer adjudication that divulges into political matters.\(^{99}\)

Two significant Articles of the Statue of the ICJ are **Article 38** and **41**.\(^{100}\) These Articles are important because **Article 38** outlines the sources of international law used by the Court in deciding the dispute and **Article 41** states the Court’s power to indicate provisional measures should it be necessary.\(^{101}\) Furthermore, if the ICJ adjudicated the disputes under the IWT between India and Pakistan, the Court would have dealt with all the disputes in an orderly manner as they will be dealt with in the paragraphs below:

The first dispute that arose under the IWT was the ‘**Salal Hydroelectric Project**’.\(^{102}\) This project as mentioned before is located on River Chenab and is a run-of-river project.\(^{103}\) Pakistan objected to the design of this project that included ‘six low-level outlets and overall height of spillway gates of 40 ft’ in

\(^{96}\) Buergenthal and Murphy, *Public International Law in a Nutshell*, Ch. 4, at Pg. 78. (4th ed. 2006).

\(^{97}\) Id. at 79.

\(^{98}\) *Statute of the International Court of Justice*, Article 36 (1) *The jurisdiction of the Court comprises all cases which the parties refer to it and all matters specially provided for in the Charter of the United Nations or in treaties and conventions in force*.


\(^{100}\) *Statute of the International Court of Justice*.

\(^{101}\) Id. at Article 38 and 41.

\(^{102}\) Iqtidar Siddiqui, *supra* at 95.

\(^{103}\) Azhar Ahmad, *supra* at 4.
clear violation of the treaty.\textsuperscript{104} This dispute was amicably resolved bilaterally by way of an agreement between both the countries in 1978 that included reducing the height of the spillway gates from 40 to 30 ft.\textsuperscript{105} The way ICJ would have decided this dispute would have been in light of the IWT provisions. Even the Court would accept upon interpreting the treaty provisions that India is allowed to store and construct storage works in restrictive conditions, however, it cannot go beyond them.\textsuperscript{106} After hearing both states’ arguments and examining evidence thoroughly, the Court having regard to the treaty provisions would have asked India to reduce the height of spillway gates and alter the design in a way that does not violate the treaty provisions. Additionally, the Court in my view would have ordered the height of the spillway gates to be reduced to at least 20 ft instead of 30 ft agreed between Pakistan and India bilaterally.

The next dispute concerns the ‘\textbf{Wullar Barrage Project}', also known as ‘Tulbul Navigation’.\textsuperscript{107} Wullar Lake is the largest fresh water lake based on River Jhelum and India has planned to construct a barrage at the outfall of the lake into the river.\textsuperscript{108} The construction started in 1984 devoid of any information to Pakistan as required under IWT in spite of the size of the project.\textsuperscript{109} Additionally, under IWT India is not allowed to construct any storage works of the main River Jhelum except its tributaries very restrictively.\textsuperscript{110} Pakistan only became aware of this project in February 1985.\textsuperscript{111} Pakistan contends that India would be able to gain regulation control of the Wullar Lake by creating supplementary storage volume that would aide in amassing water for up to 6 months during wet season.\textsuperscript{112} The importance of this project is that it could potentially destroy the intact ‘triple canal project’ of Pakistan.\textsuperscript{113}

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{104} Id.
\item \textsuperscript{105} Iqtidar Siddiqui, \textit{supra}, at 96.
\item \textsuperscript{106} IWT 1960, Article III (4).
\item \textsuperscript{107} Iqtidar Siddiqui, \textit{supra} at 96.
\item \textsuperscript{108} Id.
\item \textsuperscript{109} Id. at 97.
\item \textsuperscript{110} Id.
\item \textsuperscript{111} Iqtidar Siddiqui, \textit{supra} at 96.
\item \textsuperscript{112} Azhar Ahmad, \textit{supra} at 5.
\item \textsuperscript{113} Id.
\end{itemize}
\end{footnotesize}
Furthermore, construction works were suspended by India in 1987 after Pakistan strongly protested against the features of this project as according to the treaty, India is only allowed to construct such projects on Main Chenab and Jhelum Rivers with a total storage capacity of not more than 10,000 acre ft.\(^\text{114}\) However, the current storage capacity of the barrage is 30,000-acre ft, which exceeds the allowance.\(^\text{115}\) The PIC tried to resolve this issue as required under Article IX (1) of IWT but it was unsuccessful in doing so and both India and Pakistan are endeavouring to resolve the dispute but there has been no success so far.\(^\text{116}\)

India’s contention on this issue is that it is allowed and has the ‘technical endorsement’ in light of Article III (4) of IWT to go ahead with this project.\(^\text{117}\) Though, it is generally prohibited to build such projects under the treaty but India is of the view that it can construct this project as an exception and under certain conditions having regards to technical specifications as outlined in Annex D and E of the treaty for navigational use.\(^\text{118}\) Moreover, the project will prove to be beneficial to Pakistan as the supplies of River Jhelum would regulate to Pakistan's Mangla reservoir.\(^\text{119}\)

On the other hand, Pakistan’s position is that India violates the treaty by building this project, as under Article II (inter alia) the distribution of water is clearly stated and India is not allowed to block or store water of River Jhelum.\(^\text{120}\) Pakistan further contends that firstly, the treaty does not permit the purpose behind this project and secondly, various improved links are available as an enhanced substitute to navigating the river, therefore, the need to construct this barrage is not vital.\(^\text{121}\)

\(^{114}\) IWT 1960, Annexure E, sub-para 8 (h).
\(^{115}\) Azhar Ahmad, supra at 5.
\(^{116}\) Iqtidar Siddiqui, supra at 97.
\(^{117}\) Id. at 98.
\(^{118}\) Id.
\(^{119}\) Id.
\(^{120}\) Id.
\(^{121}\) Id. at 99.
ICJ after having heard both India and Pakistan’s contentions, in the first instance, would have indicated ‘provisional measures’ under Article 41 of its Statute in preserving Pakistan’s rights vested under Article III (4) of the treaty by ordering India to suspend any works on the project until the Court gives its final decision. In its final decision, after considering both India and Pakistan’s position on the issue in light of the treaty provisions it would have decided that India is in violation of IWT by building Wullar barrage on River Jhelum as it is clearly not allowed to do so under the treaty.122

Although, India puts forward a reasonable argument in constructing this project solely for navigational use but according to the principles of customary international law and Vienna Convention on the Law of Treaties 1980, as ICJ will apply these under Article 38 of its Statute would conclude that India has breached the treaty and should discontinue this project. The Court would reach this decision after considering the principle of pacta sunt servanda123 and treaty interpretation under Article 31(1) of VCLT 1980 that stipulates ‘a treaty shall be interpreted in good faith in accordance with the ordinary meaning to be given to the terms of the treaty on their context and in the light of its object and purpose’. Therefore, ICJ in applying these principles that the parties to a treaty ought to honour their promises and ordinary meaning is to be given when interpreting treaties having due regard to the object and purpose of the treat would aide in reaching the above-mentioned decision. This is because in giving ordinary meaning to Article III (4) of the treaty and its annexes, India’s argument of exception is void as the object and purpose of IWT is regulating the flow of water in the Indus Basin apart from sharing international water recourse in an equitable manner.

The other dispute that has arisen is ‘Baglihar Hydroelectric Project’. This is a run-of-river project located on River Chenab.124 The approval was given in 1996

---

122 IWT 1960, Article III, Para 4.
123 VCLT 1980, Article 26: ‘every treaty in force is binding upon the parties to it and must be performed by them in good faith’.
124 Azhar Ahmad, supra at 4.
and the construction started in 1999. Since October 2008, this project has been in operation. The features of this development include a high concrete gravity dam with a gross storage and power generation capacity. India permitted Pakistan’s team to inspect the site in 2003. Objections related to the design of under-sluice type gated spillway, location of power intake, excessive free board, and pond level in violation of IWT were raised by Pakistan. The PIC was unable to resolve the contentions raised by Pakistan and the matter was referred to the World Bank for the appointment of a neutral expert. The World Bank then appointed Mr. Raymond Lafitte, after consulting both India and Pakistan, who gave his Final Determination in 2007.

The neutral expert’s determination upheld disputation regarding the plant’s design that were raised by Pakistan and it established that India’s design did not conform to the design criteria of all the features of the dam. In essence, Mr. Lafitte recommended India to reduce free board from 4.5 to 3 metres but did not ask to change the spillway design relying on international practice. He also concluded that ‘the calculation methodology used by India was not in conformity with the object and purpose of the treaty’. Shortly after the determination, India obstructed River Chenab’s flow whilst filing Baglihar reservoir in 10 days instead of spreading it out over the span of 60 days, thereby, violating the treaty provisions. This resulted in loss of water flow received by Pakistan as much as 0.2 million acres foot of water that greatly affected its agriculture sector even after numerous requests to India to provide Pakistan with the filling schedule. The issue remains unresolved so far as Pakistan has sought compensation from India for the lost waters and India accepts Pakistan’s claim but it is inflexible.

---

125 Id.
126 Iqtidar Siddiqui, supra at 99.
127 Ahmad, supra at 4.
128 IWT 1960, Para 8, Annexure D.
129 Iqtidar Siddiqui, supra at 101.
130 Id.
131 Id.
132 Ahmad, supra at 5.
133 Siddiqui, supra at 102.
134 Ahmad, supra at 5.
135 Id.
with regards to this matter.\textsuperscript{136} Therefore, Pakistan has demanded monitoring of River Chenab’s water flow by setting up telemetry device.\textsuperscript{137}

ICJ would have addressed this dispute by siding with the neutral experts as this an extremely technical issue and would require an expert’s opinion. The Court would have handed down ‘provisional measures’ to prevent any construction work from proceeding further until the final decision.\textsuperscript{138} It would have then, after starting the case, appointed an expert who would have conducted his determination in exactly the same manner as done by Mr. Lafitte.\textsuperscript{139} The Court would then have given its final decision on the basis of expert determination and would have asked India to bring the project’s design into conformity with the objective and purpose of the IWT, the equitable use of this watercourse.

The second issue that would have been addressed with regards to the same dispute would be the filling of Baglihar reservoir that caused a reduction in River Chenab’s flow during significant agriculture cultivation months.\textsuperscript{140} The Court, after taking India’s circumstances when filling the reservoir into consideration, would have reached the conclusion that there was no immediate need or any urgency to fill the dam. This act has caused great consequences to Pakistani agriculture sector as due to the reduction in the flow of water, the crops on more than 2.5 million acres were damaged.\textsuperscript{141} Though, India accepts Pakistan’s claim of reduction in the water flow of River Chenab\textsuperscript{142} but it has done nothing to compensate Pakistan of this loss. Pakistan contends that India should recompense it either by paying damages or by returning the same amount of water that it blocked to fill the dam.\textsuperscript{143} Additionally, Pakistan demands that a telemetry device should be set up to monitor the water flow of River Chenab.\textsuperscript{144} The Court on this point would have ordered India to compensate Pakistan by

\textsuperscript{136} Siddiqui, supra at 103.

\textsuperscript{137} Ahmad, supra at 5.

\textsuperscript{138} Statute of ICJ, Article 41.

\textsuperscript{139} Professor Raymond Lafitte, Swiss Federal Institute of Technology, (ICOLD).

\textsuperscript{140} Siddiqui, supra at 102-103.

\textsuperscript{141} Id. at 104.

\textsuperscript{142} Id.

\textsuperscript{143} Ahmad, supra at 5.

\textsuperscript{144} Id.
returning the same amount of water and for Pakistan to store this water for agriculture purposes and not let it flow into the Arabian Sea. In my view, the Court would have ordered this particular sort of compensation as the real issue is of water and not damages. Furthermore, the Court would have ordered installation of telemetry device to monitor the flow of water to avoid any confusion or violation of this kind in the future. In reaching this decision, the Court would have had regard to the treaty provisions as well as the ‘no harm to the lower riparian’ principle.\(^{145}\)

The other alleged violation of IWT is ‘Kishangana Hydroelectric Project’. This is being constructed by India on the Neelum River.\(^{146}\) India started the construction of a tunnel on Neelum River to divert the waters into the Wullar Lake.\(^{147}\) The project is designed to divert the flow of this dam into Wullar Lake through a 21.6 KM tunnel after generating 330 MW of power. Under the treaty, India is allowed legal construction of storage works for up to 0.75 million acres feet (MAF) on the Jhelum River tributaries.\(^{148}\)

The ICJ would essentially be asked to rule on the legality of this project and whether it should be allowed to proceed in light of the treaty provision and customary international law?

First of all the objections raised by Pakistan should be stated and analysed. Pakistan contends that the project contravenes Para 10 of Annexure E of IWT 1960.\(^{149}\) The other contention is that it is not allowed under the treaty to divert the flow of one tributary of Jhelum River into another of its tributary.\(^{150}\) Furthermore, the project’s design parameters are not in conformity with that

---


\(^{146}\) Azhar Ahmad, Indus Water Treaty A Dispassionate Analysis, Policy Perspectives, (2011).

\(^{147}\) Id.

\(^{148}\) IWT 1960, Annexure E, para 7.

\(^{149}\) “Notwithstanding the provisions of Paragraph 7, any storage work to be constructed on a tributary of the Jhelum on which Pakistan has any agricultural use or hydro-electric use shall be so designed and operated as not to adversely affect the then existing Agricultural Use or hydro-electric use on that Tributary”.

\(^{150}\) IWT 1960, Annexure E.
laid down in the treaty.\textsuperscript{151} Pakistan also puts forward a detailed explanation as to how this project will affect its very own hydroelectric power project and the agriculture use in the area in addition to the direct bearing it will have on socio-economic life and ecological aspects in the area down stream of the plant.\textsuperscript{152} Additionally, Pakistan has also started its own project ‘Neelum-Jhelum Hydroelectric Project’ in Pakistan administered Kashmir. This project is proposed to generate 969 MW of power and due to operate in 2017, which is a year after Kishangana is due to begin its operation.\textsuperscript{153} However, Pakistan contends that Indian project’s proposed design will result into 13.4 percent reduction in annual average inflows of Neelum-Jhelum River dam site and a power reduction of 10.2 percent.\textsuperscript{154} Nevertheless, Pakistan’s main objections are with regards to the design parameters of the project and diverting the flow of one tributary to another because it believes that its own hydroelectric project will be deprived of some serious discharges resulting the production capacity to be reduced from 11 to 16 percent.\textsuperscript{155}

In India’s response to these contentions, the Indian Commissioner stated that a revised design of the project from Storage Work to Run-of-River project is under consideration,\textsuperscript{156} that being allowed under the treaty and reduction of the dam’s height by 40 metres was in progress.\textsuperscript{157} The revised project information was also provided to Pakistan as required under the treaty.\textsuperscript{158} However, diversion of water along with other characteristics of the project remains unaltered. IWT allows India to divert water of a tributary of Jhelum River to another tributary in relation to Run-of-River projects; however, it is subject to taking care of Pakistan’s existing Agricultural and hydroelectric uses.\textsuperscript{159}

\begin{footnotesize}
\begin{enumerate}
\item[151] Id. Para 11 (a), (c), (e), (f), and (g).
\item[153] Dr. Shabir Coudhry, \textit{Neelum-Jhelum Hydroelectric Project – An Environment Disaster}, (2010), available at: \url{http://www.countercurrents.org/choudhry260610.htm}.
\item[154] Id. Footnote 7.
\item[155] Azhar Ahmad, \textit{supra} at 6.
\item[156] IWT 1960, Annexure E, Para 15A.
\item[158] IWT 1960, Annexure D, Appendix II.
\item[159] Id. Footnote 156.
\end{enumerate}
\end{footnotesize}
As the Permanent Indus Commission declared the difference as a dispute, Pakistan referred it to PCA for arbitration and as a result interim measures were granted in Pakistan's favour. The interim measures essentially held that 'India may not construct any other permanent works on or above the riverbed that may inhibit the restoration of the full flow of that river to its natural channel'.\(^{160}\) The dispute is still on going pending final award by PCA.\(^{161}\)

ICJ would have decided this contentious issue by giving 'provisional measures'\(^{162}\) just as PCA handed down interim measures in the first instance but ICJ's measures would have been strict in that India would not have been allowed to continue any type of works on this project, let alone permanent works. In its final decision, the Court would have ordered India to revise the design project that is in conformity with the treaty especially addressing the issue of water diversion. Upon interpreting Article III, IV and Paragraph 15 (iii) of Appendix D of the treaty and giving these provisions their ordinary meaning, the Court would have held that diverting the water of one tributary into another is not permitted in light of these provisions.

Furthermore, the Indian project unduly interferes with the Pakistani hydroelectric project besides having adverse effects on the agriculture and ecological uses of Pakistan; therefore, it should revise the design parameters on the basis of no harm to the lower riparian principle of customary international law.\(^{163}\) Moreover, when India has revised the plan it must consult the Court and get its approval before resuming any constructing works.

The Court upon assuming jurisdiction on Dul Hasti Hydroelectric Plant issue, which is a 390 MW 'high concrete gravity dam' upstream of Baglihar hydroelectric project under construction on River Chenab\(^{164}\) would have decided


\(^{161}\) Id.

\(^{162}\) Statute of ICJ, Article 41.


\(^{164}\) Azhar Ahmad, supra at 6.
the case on the principles of customary international law as the treaty provisions are not in actual fact being violated. Though the adverse effects of this particular project on Pakistan are not severe but it is vital to prevent India from building projects because of Pakistan's concern that it might expand on the project in the future. The Court would have ordered India to provide detailed design parameters of this project to Pakistan and if Pakistan after examining them does not raise any solid objection (s), then India may well proceed with this project in accordance with IWT. However, if Pakistan raises objection (s) then the dispute should be referred to ICJ for the matter to be adjudicated.

The other dispute that could be referred to the ICJ is 'Bursar Dam'. The project is currently under construction and is being built on both River Jhelum and Chenab flowing through Indian Occupied Kashmir.165 This is thought to be the major project that would generate 1020 MW of electricity.166 This dam is considered to be a storage facility anticipated to be completed over the span of 6-7 years that would regulate the water flow to downstream projects.167 Moreover, the current storage capacity of 2.2 MAF is considered beyond the limits permitted under the treaty.168 Not only does this particular project violate the treaty, but it also violates international environmental conventions. This project does not only cause water scarcity in Pakistan but also aides to be a contributing factor for melting of Himalayan glaciers.169

ICJ after analysing would have decided that this project is primarily a storage facility, which is a flagrant breach of the treaty as projects of such nature are explicitly prohibited under the treaty.170 The Court would have ordered India to either change the design of this project to bring it in line with the treaty provisions or discontinue its plans to build this project. This particular dam causes water scarcity in Pakistan, thereby causing harm to it in contravention of

---

165 Id.
166 Id.
167 Id.
168 Id.
169 Id.
170 IWT 1960, Article III (4).
international environmental law.\textsuperscript{171} India has an obligation under the treaty and international environmental law to not cause harm to its lower riparian and make use of the water in an equitable manner.\textsuperscript{172}

The other contentious issue that ICJ would rule on is the ‘\textbf{Nimoo Buzgo Hydropower Project’}. This is a run-of-river hydroelectric project on River Indus with an installed capacity of 45 MW.\textsuperscript{173} Pakistan only became aware of this in 2002 and requested India for further information relating to this project but until December 2006, India failed to respond despite numerous reminders.\textsuperscript{174} Since Pakistan has inspected the site and has learnt that the project is now 80 percent complete and would be inaugurated in August 2012\textsuperscript{175} there is little that the Court could do at this point as most of the construction works have been completed. However, what the Court could and would do is order India, in light of the treaty provisions, to compensate Pakistan by way of damages, as the Court may deem appropriate. This is because the nature of this project is run-of-river and these kinds of projects are permitted under the treaty. Furthermore, the Court will be mindful of the fact that India did flagrantly violate the treaty by not informing, seeking approval and/or getting the project’s design scrutinised prior to starting construction works.\textsuperscript{176} Therefore, the Court will arrive at the conclusion that India should compensate Pakistan for breaching the treaty by not informing or getting its project’s design scrutinised as required.

\textbf{Compliance With The ICJ Decisions}

The whole adjudication process boils down to compliance with the decisions handed down by any Court. Any state that brings a claim against another state and the Court gives a binding judgment would not be pleased if the state against whom the case was brought does not comply with the Court’s order. More relevant to the context of this paper, ICJ judgments in contentious cases are

\textsuperscript{171} The UN Convention of the Law of the Non-Navigational Uses of International Watercourses 1997, Article 5 and 7.
\textsuperscript{172} Id.
\textsuperscript{174} Id.
\textsuperscript{175} Id.
\textsuperscript{176} IWT 1960.
binding on the parties as stated in Article 59 of the Court’s Statute.\textsuperscript{177} The judgments are also considered to be ‘final and without appeal’ as affirmed in Article 60 of the Statute of ICJ.\textsuperscript{178} However, a review of the decision is possible only if it satisfies various conditions.\textsuperscript{179}

The ICJ’s Statute does not explicitly state how the Court’s decisions are to be enforced, however, States have generally been compliant with the judgments except in few occasions.\textsuperscript{180} On the contrary, the UN Charter deals with the issue of ICJ decisions’ compliance and it states ‘each UN member state undertakes to comply with the decision of the International Court of Justice in any case to which it is a party’.\textsuperscript{181} This means that any member state that does not comply with the decision of ICJ would thus violate the UN Charter\textsuperscript{182} and Pakistan and India are member states of the UN.\textsuperscript{183} Furthermore, the redress to a party in a case that complies with the decision against the non-compliant party is that it can appeal the non-compliance to the UN Security Council.\textsuperscript{184} The Security Council ‘may, if it deems necessary, make recommendations or decide upon measures to be taken to give effect to the judgment’.\textsuperscript{185}

In certain instances, if a state falls short of complying with the ICJ’s decision, this could imply enforcement measures.\textsuperscript{186} The Security Council can take these enforcement measures against the non-compliant state by asking its permanent member to use their veto power or in the alternative impose sanctions. This action can only be taken against the non-compliant member if the permanent members are ready to cooperate.\textsuperscript{187} The Security Council may respond in one of the following ways. Firstly, take no action at all; secondly, it may ask the non-compliant country to comply with the decision, thereby adding importance to the

\begin{footnotesize}
\begin{enumerate}
\item[177] Buergenthal and Murphy, \textit{supra} at 90.
\item[178] Id.
\item[179] Statute of ICJ, Article 61.
\item[180] Buergenthal and Murphy, \textit{supra} at 91.
\item[181] UN Charter, Article 94 (1).
\item[182] Buergenthal and Murphy, \textit{supra} at 91.
\item[184] UN Charter, Article 94 (2).
\item[185] Id. Footnote 181.
\item[186] UN Charter, Articles 39, 41 and 42.
\item[187] Buergenthal and Murphy, \textit{supra} at 91.
\end{enumerate}
\end{footnotesize}
judgment; thirdly, it may make recommendations to facilitate implementation of the judgment that could include intercession through economic or diplomatic sanctions or use of force in worst case scenario; lastly, it may ask member states to exercise political pressure upon the non-compliant state.\textsuperscript{188} However, the Security Council so far has not implemented measures on the basis of Article \textbf{94(1)} of the UN Charter and has hardly ever interceded in the enforcement of an ICJ decision.\textsuperscript{189}

According to a study by Dr Schulte on compliance record of the ICJ decisions from 1946 to 2003, she concluded that out of 27 distinct cases as of the end of 2003 that reached a judgment on the merits, a ‘generally satisfactory compliance record for judgments’ was achieved.\textsuperscript{190} Therefore, the statistics show that there is a good track record of compliance with the ICJ decisions and in my view this is due to the fact that ICJ is the principal judicial organ of the UN\textsuperscript{191} and member states generally avoid running into problems with the UN.

**Conclusion**

Whilst some scholars argue that ICJ is ‘limited in its ability to effectively resolve cases involving environmental issues’ and put forward the following reasons in support of this argument.\textsuperscript{192} Firstly, they argue that ICJ is the venue of last resort for nations involved in interstate disputes. Secondly, the developing nations concern that the Court lacks diversity and has a set western standard of adjudicatory process. Thirdly, the Court does not have the competence to deliberate all possible disputes over which it could presume jurisdiction. Lastly, the proceedings in this Court are costly, slow and often the result rendered does not impart an adequate remedy.\textsuperscript{193} Conversely, I am of the view that ICJ is the most suitable forum for interstate disputes because it can impart with decisions

\textsuperscript{189} Id.
\textsuperscript{190} Constanze Schulte, *Compliance with Decisions of the International Court of Justice*, at 89 and 403, (2004).
\textsuperscript{191} Buergenthal and Murphy, *supra* at 78.
\textsuperscript{192} Anna Spain, *supra* at 363.
\textsuperscript{193} Id. at 364.
that are acceptable to and welcomed by the parties as it has done so far.\textsuperscript{194} This method of dispute resolution could work best for both India and Pakistan as they have already tried various methods that have proved not to be very fruitful in terms of compliance. The fact that the compliant state has recourse to refer non-compliance to the UN Security Council, which could then take extreme measure to ensure compliance with the ICJ decisions, is an added benefit for states to refer their disputes to the ICJ for adjudication.

In this era of climate change and global warming, both Pakistan and India will have to make use of the Indus Basin efficiently and effectively. Pakistan, in particular, should store water by building dams and reservoirs to meet the increased agriculture and hydroelectric needs, instead of wasting the resource by allowing it to flow into the Arabian Sea. India, on the other hand, has to be mindful of its current hydroelectric projects strategy so as to not cause harm to its lower riparian. The IWT acknowledges that both India and Pakistan have an interest in the optimum development of the rivers, and to that effect provides for cooperation and collaboration between the two countries.\textsuperscript{195} The underlying problem is that the IWT did not and could not have foreseen the current hydroelectric needs of both India and Pakistan at the time it was signed; therefore, it fails to address this issue accordingly. As a result of this, I propose that the IWT must be revisited and reconsidered to address the needs of both states whilst preserving the ecosystem of the Indus Basin. Additionally, the revised version of the treaty must specifically modify Article IX of the treaty so as to include ICJ as the main forum for dispute resolution. Otherwise, there is a fear that India’s repeated violations of IWT may lead to South Asia’s first water war between two neighbours that are nuclear powers.

In conclusion, though it is true that proceedings in ICJ are slow, costly, and considered to be the last resort but such is the case with other venues too as all the dispute resolution forums in the IWT are also slow and costly. Nevertheless, the fact that this forum has never been tried and most of the others have been

\textsuperscript{194} Example: \textit{Gabcikovo-Nagymaros Project Case} (Hungary v. Slovakia).

\textsuperscript{195} R.K.Ahora, \textit{supra} at 13.
raises a probability, even if it is slight, that there could be some positive progress on these disputes in the ICJ towards the right direction even if the disputes are not resolved in full. I support this view because Pakistan and India have reached a point where no other dispute resolution mechanism seems to resolve the issue satisfactorily and the only ray of hope is in approaching the ICJ to adjudicate these disputes fairly and peacefully.

AQEEL MALIK
Institutional Reform of Competition Law Authorities: Accounting for Psychological and Economic Factors that Influence Behaviour

SAM PRICE

The European Commission (henceforth ‘the Commission’) and National Competition Authorities (NCAs) conduct investigations into potential breaches of Competition Law and also decide whether there has been a breach. Some scholars have maintained that in the absence of an independent judge, or other impartial adjudicator, such procedure violates Article 6(1)(c) of the European Convention on Human Rights. Wils has highlighted certain economic and psychological motivations that may impact upon the procedural fairness of antitrust investigations. This essay utilises the framework of Wils to scrutinise the current system, which merges prosecutorial and adjudicative functions, and the extent to which reform could alleviate these issues. It is suggested that psychological and economic factors, as well as restrictions due to the competence of potential adjudicators, means that the current system does not suffer a deficit in regard to certain key factors compared to alternatives. These factors include independence, expertise and due process. While this essay does not argue against reform per se, it seeks to examine how the deficits highlighted within the current system may be equally present in alternative schemes.

Section 1 of this essay presents a context to the debate. It first assesses Competition Law rules as substantive provisions noting the preference for effect-based methodologies which may require a different procedural approach to per se rules. The majority of section 1 is concerned with the institutional layout and the powers of the institutions in enforcing Competition Law as well as with

2 Ibid.
3 For this essay investigations may be used as shorthand meaning the investigation, prosecution and decision about the guilt of an undertaking or association in relation to an antitrust violation.
alternative structures. Section 2 utilises psychological and economic literature to suggest the types of unwanted behaviour that may occur within the current system and which may also occur within alternate systems arguing that alternatives are only superficially preferable. When the checks and balances within the systems are evaluated, preventing the unwanted behaviour is not as simple as requiring a separate adjudicative body. This is especially so since the European system will always have some form of appeal to the Courts due to the operation of the Treaty on the Functioning of the European Union (TFEU).

Section 1: The Current System and Potential for Reform

Since the procedural structure of Competition Law institutions is designed to allow the optimum application of the substantive rules, a point on the substantive rules must be noted. There has been significant reform in enforcement decisions and procedures relating to Article 102 TFEU with authorities willing to utilise economic-based methodologies rather than relying on per se rules.\(^5\) Recommendations for this approach were made in the European Commission consultation paper on the matter which favoured an effects-based approach which has consistently been the preferred approach in relation to Article 101 and merger control.\(^6\) The implication of this approach being dominant in Competition Law is that certain institutions that may be otherwise preferable are unsuitable; for instance, Ezrachi states "it is generally accepted that the majority of national courts have limited capacity when it comes to the analysis and deliberation of complex economic concepts and the assessment of consumer harm".\(^7\)

---

\(^7\) Ezrachi, ‘The interplay between the economic approach to Article 82 and private enforcement’ (2008) G.C.L.R. 148, 151.
Moving on to the structure of public enforcement, each EU Member State has NCAs and Competition Laws, though as with all EU matters, EU law is supreme; these varying nation agencies work as part of a wider grid – the European Competition Network (ECN) – which also incorporates the work of the European Commission. The Commission and NCAs investigate and adjudicate in Competition Law cases determining the facts and whether national or European law has been breached. The current structure follows a period of decentralisation in which Member States’ agencies were further empowered. Following Regulation 1/2003 NCAs can apply Article 101(3) as well as 101(1), 101(2) and 102 TFEU. National Courts also share competency.

On the UK level, the OFT and Competition Commission are the bodies designated under Article 35 Regulation 1/2003 as the bodies who have such competency. At present, the OFT investigate Competition Law infringements and criminal offences under Enterprise Act 2002 as well as phase 1 mergers. The Competition Commission investigate phase 2 mergers referred from the OFT and conduct market enquiries referred by the OFT. The UK institutions are currently undergoing a reform which will see the OFT and Competition Commission merged to form the Competition and Markets Authority by 2014. At the European level, Recital 15 of Regulation 1/2003 requires the Commission and NCAs to work as a collaborative whole to investigate, prosecute and adjudicate against undertakings and associations who contravene Competition Law. Members of the ECN communicate via a secure intranet. Recital 21 and 22 involve national courts and require harmonious application of Competition Law respectively. The Commission has also published notices on cooperation. NCAs have the similar powers to the Commission in relation to decision making about infringements, imposing fines, accepting commitments, etc. However, NCAs

---

11 See OJ [2004] C 101/43 Notice on cooperation within the network of competition authorities and OJ [2004] C 101/54 Notice on the cooperation between the Commission and courts of the EU Member States in the application of Articles [101 and 102 TFEU].
cannot determine that EU Competition Law rules are not infringed, only that there are no grounds for action.\textsuperscript{12} The setup discussed clearly indicates that the ECN is one large bureaucracy.

The ECN, particularly the Commission, enjoys certain powers conferred by Regulation 1/2003 including adjudicative powers, investigatory powers and powers to impose penalties (these are contained within Chapters III, V and VI respectively). Article 7(1) grants to the Commission the power to require the cessation of behaviour which it decides, under prescribed decision-making procedure, violates Competition Law and allows the Commission to impose behavioural or structural remedies. It is this article which those opposing joint prosecutorial and adjudicative functions find problematic. The Commission can request all necessary information to carry out its duties.\textsuperscript{13} The Commission can make a simple request for information or can require information by decision. A simple request under Article 18(2) requires that the Commission state the legal basis and purpose of the request as well as specifying the desired information and a time limit. While there is no duty to comply with a request, if a firm provides incorrect or misleading information it may be penalised under Article 23, though the Commission must inform it of such a possibility when making a request. Requiring information by decision is covered by Article 18(3) and non-compliance is punishable under Article 23 though undertakings have the right to have the decision reviewed by the General Court regarding purpose and proportionality.\textsuperscript{14} Articles 21 and 22 respectively allow the Commission to search business premises and other premises including individuals’ homes. The former encompasses ‘all necessary inspections’ and permits surprise investigations (‘dawn raids’) while the latter requires judicial sanction. While the powers discussed in this paragraph may be far reaching, they have limitations and the possibility of review and procedural requirements are designed to ensure that the Commission acts in a transparent and accountable way.

\textsuperscript{12} Case C-375/09 Prezes Urzedu Ochrony Konkurencji I Konsumentow v Tele2 Polska sp. Z o. o. now Netia SA [2011] ECR 1-000.
\textsuperscript{14} On proportionality see ibid, SEP v Commission.
Since the investigation, prosecution and adjudication of a matter are carried out by the same bodies acting collaboratively, the right to appeal is particularly important. Within the UK, appeals on the merits of cases where the decisions are appealable go to the Competition Appeal Tribunal (CAT) with the Administrative Court conducting judicial review in relation to other matters. Appeals on points of law can go before the Court of Appeal and the Supreme Court. CAT, the Court of Appeal and the Supreme Court can also request preliminary rulings under Article 267 TFEU. On the European level, applications under Articles 261, 263 or 265 TFEU can be brought before the General Court and the Court of Justice hears appeals on points of law.

There are alternatives which are briefly laid out here for the purpose of comparison. The US has two systems. The Federal Trade Commission (FTC) has investigatory powers and limited adjudicative powers. The FTC can reach settlements or dismiss investigations; if an infringement is found it can bring litigation against the defendant with the FTC’s complaint council and the defence conducting a trial before an administrative judge. The FTC can hear appeals from the administrative judge from either ground. The defendant, if prejudiced can then appeal to the Federal Court of Appeal and, potentially though unlikely, the Supreme Court. The FTC is therefore prosecutor and adjudicator upon appeal. The Department of Justice, another US body, has only investigatory and prosecutorial powers. Any case is brought before a Federal district court with right of appeal before the Federal Court of Appeal and, potentially, the Supreme Court. In the EU these systems could be mimicked with the Commission having a department of impartial adjudicators trained to deal with antitrust cases (henceforth ‘the FTC model’) or with cases being brought before national courts or the General Court depending on whether an NCA or the Commission brings the case (henceforth ‘the DOJ model’).

15 Whish, Competition Law (7th, OUP, 2012) 80.
16 Ibid 79.
17 Wils (n 1) 208.
Section 2: Unwanted Behaviour and Reform’s (Limited) Potential to Alleviate It

Various bias and rent seeking behaviour can occur within procedures which distort the potential for the substantive aims of provisions to be achieved. This section lays out the psychological and economic factors behind this behaviour and how such behaviour may manifest. For each of these it is considered where these behaviours may arise in the current system and the potential of alternatives in limiting such occurrences. The psychological and economic factors considered are: belief bias, hindsight bias, group-perception related behaviour and rent seeking and similar bureaucratic behaviour. The competence of individuals is also discussed to the extent that it has implications upon the right to a fair trial – certainly errors can be categorised as unwanted behaviour. It should also be reiterated that the proceeding analysis does not conclusively advocate one institutional layout as preferable; this would require further economic and legal analysis; this section simply illustrates certain behavioural factors that must be considered.

Belief bias and closely associated psychological phenomena occur where an individual’s objective assessment is distorted by their personal views. This belief bias can affect the extent to which individuals accept the validity of statements or evidence presented to them.\(^\text{18}\) Confirmation bias also exists which is "an inclination to retain, or a disinclination to abandon, a currently favoured hypothesis";\(^\text{19}\) this links closely to a related bias known as ‘positive hypothesis testing’ where individuals look to confirm what they expect.\(^\text{20}\) Another type of bias is known as ‘contrasting effect’ occurs where individuals exaggerate the difference their own behaviour and that of others.\(^\text{21}\) This frequently occurs

\(^{18}\) This has been illustrated repeatedly in syllogistic tests where participants’ acceptance of logical statements is profoundly affected by political, moral, etc views – see Evans, Barston and Pollard, ‘On the conflict between logic and belief in syllogistic reasoning’ (1983) 11 Memory and Cognition 295.


\(^{20}\) Ibid.

where the other individuals are held in low esteem.\textsuperscript{22} These biases can occur within antitrust investigation initiated by the ECN. As Wils highlights, the extent to which these biases occur in phase 1 merger investigations may be limited as all potential mergers must be notified to the Commission where there is a community aspect\textsuperscript{23} and most are unproblematic though biases may occur in phase 2 investigations, the trigger for which is potential for the merger to contravene Competition Law.\textsuperscript{24} Conversely, investigations for breaches of Article 101 and 102 initiate on the basis of complaints or suspicions that a Competition Law breach is occurring. In these types of investigations, as with phase 2 mergers, ECN workers are likely to believe that the ECN only initiates investigations with good reason and that in all likelihood a breach has occurred. In this case they will be more prone to find evidence supporting the notion that a breach has occurred and they will be more likely to ignore contrary information. This is particularly worrying where these individuals will be providing complicated economic analysis which may be open to some interpretation.

Hindsight bias occurs when the knowledge of an outcome “influences the judgements we make for a naive other or a naive prior-self”.\textsuperscript{25} Awareness of the eventual outcome leads to distorted perception where individuals feel they ‘knew it all along’.\textsuperscript{26} This has the potential to affect the way that superiors, observers or colleagues recount the actions of ECN employees. It may also affect employees’ self-perception.\textsuperscript{27} For the latter, this can result in ‘cognitive dissonance’ where individuals feel that an action was not justified due to hindsight bias thereby causing self-doubt.\textsuperscript{28} Individuals actively try to avoid this self-doubt and so they will avoid discoveries that reveal their actions ‘should not have been taken’. Individuals, subconsciously aware of hindsight bias – that is to

\textsuperscript{22} Dawes, Singer and Lemons, ‘An experimental analysis of the contrast effect and its implications for intergroup communication and the indirect assessment of attitudes’ (1972) 21 J.P.S.P. 281, 281.
\textsuperscript{23} Regulation 139/2004 OJ [2004] L 24/1 under section 4 (the EU Merger Regulation).
\textsuperscript{24} Wils (n 1) 217.
\textsuperscript{25} Bernstein et al, ‘Hindsight Bias and Developing Theories of Mind’ (2007) 78 Child Development 1374, 1374.
\textsuperscript{27} Wils (n 1) 218-19.
\textsuperscript{28} Ibid.
say, knowing superiors will likely condemn actions that were justified but ultimately incorrect in a technical sense – may be inclined to ignore contradictory information and favour information that supports their actions. This is similar to confirmation bias though it is not subconscious and it links closely to rent seeking and similar behaviour. The current system may suffer this form of bias in several ways, though it also has potential to reform with organisational restructuring rather than institutional reform. For instance, bureaucrats investigating mergers may condone a phase 2 investigation which may subsequently become unnecessary in the face of more evidence. This does not mean that the original decision to move to phase 2 was incorrect though it may be perceived as such due to hindsight bias. As such, individuals investigating may be reluctant to conclude the phase 2 investigation was unnecessary and so they may try and impose superfluous commitments as a quasi-justification. This may also protect them from the potential scorn that observers or superiors may levy. This could be avoided by having different individuals conduct the second phase investigation, much like in the UK where the OFT conducts phase 1 and the Competition Commission phase 2 (though a different body would not be necessary). The issue may be more difficult for investigations into Article 101 and 102 breaches though different individuals could recommend and conduct the investigation stages.

Intergroup related psychological factors may also be influential in the functioning of an institutional system. This relates to the actions taken by individuals in one group, whether as individuals or a collective, in interaction with another group.29 Perceptions of difference between groups can lead to increased solidarity within a group and also outward hostility to other groups.30 These differences may stem from different values or cognitions or merely the perception of intergroup competition.31 Regulation 1/2003’s institutional reform may have helped mitigate jurisdictional problems stemming from such perceived competition. NCAs may still view themselves as distinct from other NCAs or from

31 Tajfel (n 29) 20-21.
the Commission. Intergroup conflict may also be apparent between the ECN and either national courts or EU level courts with the latter being more important since national courts are bound to follow Commission precedent. This tension may be reciprocated and the EU courts have discussed the need for the Commission to ensure it acts within the powers conferred by the EU treaties of which it is the interpreter.\footnote{The legality of a Commission decision can be reviewed under grounds contained in Article 263. Of particular importance in relation to this topic is the right to a fair trial. This has been at issue in Case 17/74  Transocean Marine Paint Association v Commission [1974] ECR 1063. The EU Courts have the power to review substantive and procedural factors and have been willing to do so comprehensively in the past – see Case T-210/01 General Electric v Commission [2005] ECR II-5575; [2006] 4 CMLR 686.}

The final group of factors which can instigate unwanted behaviour are rent seeking, bureaucratic failure and related economic behaviour. These can be divided into individual actions of a bureaucrat or systemic actions by the institution as a whole. Regarding the latter, bureaucracies tend to have significant informational advantages over politicians.\footnote{Thornton, \textit{Economics of Prohibition} (1st, Utah UP, 1991) citing Niskanen, \textit{Bureaucracy and Representative Government} (1st, Chicago, 1971).} This can result in a lack of regulation or intervention to counteract any potential bureaucratic failure. Bureaucracies also have a tendency to mandate the shift of resources from preferable activities to observable activities.\footnote{Lindsey, \textit{‘A theory of government enterprise’} (1976) 84 J.P.E. 1061, 1074-77.} This concern with the observable is understandable in the political sphere but can lead to failures – for instance, the Commission may be unsure whether to block or permit a merger. The merger could potentially lead to increased innovation but it may also reduce competitiveness slightly. The second of these is far more observable and so the bureaucrat is likely to be more inclined to act cautiously even if doing so is socially undesirable. The unwanted actions of individuals also have the potential to occur within the current system. Individuals may be motivated by promotions, enhanced reputation, gaining favour with superiors and so on. The bureaucracy as a collective may also want to appear active so as to justify its budget. The result of this can be increased enforcement or a desire to show such enforcement. As Wils notes, statistics on fines and other details that are published may seem impressive but people are “often unable to judge
themselves whether the prohibition decisions and the amounts of the fines imposed are indeed justified".\textsuperscript{35} A desire to show activity can also be detected in Commission speeches with Monti having previously stated that “the Commission will seek to maintain the level of activity in the future”.\textsuperscript{36} Such behaviour is not limited solely to bureaucrats and econometric study has shown that judges also exhibit such tendencies if they believe it is beneficial to their chance of obtaining a promotion or that it may reduce their job’s difficulty.\textsuperscript{37}

Having analysed the types of unwanted behaviour and their potential to materialise in the current system, one may submit that to overcome the majority of these biases would be simple: have an independent adjudicative body as in the DOJ model and FTC models. This is not necessarily so. It is possible that an independent body would suffer from confirmation bias in the same way that individuals in the current system do. If investigators perceive they are only investigating because there is good reason, an independent decision maker may believe that a case is only likely to come before them if there is good reason. Wils notes that many cases settle though the Court reviews such settlements.\textsuperscript{38} The same is true with DOJ settlements though, the independent adjudicator may generally belief, notwithstanding strong contrary evidence, that the settlement made will be preferable since it has been reached by two or more parties. The use of an independent adjudicator is unlikely to affect hindsight bias as individuals, upon the commencement of an investigation, are likely to act the same no matter who it is tried in front of since hindsight bias is a subconscious factor rather than an active motive.

The behaviour discussed may be affected by the characteristics of the independent decision maker. If the decision maker is specially trained to work in antitrust then the promotion process is likely to be different. The Judicial

\textsuperscript{35} Wils (n 1) 219.
\textsuperscript{37} Cohen, ‘The Motives of Judges Empirical Evidence from Antitrust Sentencing’ (1992) 12 I.R.L.E. 13, 29. However, this is a US study and Cohen highlights that judges may be less motivated in this way where appointment is less political.
\textsuperscript{38} Wils (n 1) 221.
Appointment Committee may lack the competence to decide who should be hired or promoted on the candidates merit in the field of antitrust due to its complex nature and so the task must fall and so the task must fall to Competition Law institutions such as the Commission or the Directorate General for Competition. In this case, the judge may have motive to act in a way that suits current competition policy and so they may exhibit rent seeking behaviour. Another issue would be who could sit as adjudicator. Ex-Commission employees may be well trained in the area but they may exhibit bias towards their old employers and confirmation bias may be stronger. Conversely, an untrained adjudicative body may be used. This also has potential problems outside of competence. Firstly, an untrained body may still transmit the subconscious bias that is present in experts.

Additionally, group related psychology may occur which limits the effectiveness of the adjudicative body at preventing bureaucratic rent seeking. The ECN may view judges as an out-group who they will act hostilely towards. An ECN worker who has their case successfully appealed may claim that this occurred because the judge does not know what they are talking about. If this feeling is systemic in the ECN, successful appeals do not affect reputation or promotion potential substantially and will therefore not prevent rent seeking. Wils argues that full judicial review could eliminate rent seeking. However, this overestimates the process and pretends judges are not subject to the same bias that bureaucrats are. Wils also concedes that judicial review is limited in that the more time elapses, the more limited the review is in preventing rent seeking.

**Conclusion**

This essay has sought to argue that reform of the current system would be superficial if it did not account for psychological and economic factors which influence behaviour. These factors and examples of how they may manifest have been provided as illustrative rather than comprehensive. Section 2 of this essay

---

39 Ibid.
40 Ibid.
also demonstrated that many biases overlap and that this can limit the
effectiveness of, for example, an independent body in reducing rent seeking and
conscious behaviour that is unwanted. Of course, it would be unwise to
recommend the optimal structure for Competition Law institutions based solely
on the preceding analysis. This is only one dimension to the debate and that
limited resources and treaty restrictions requires further fiscal and legal analysis
before any debate on reform can be seen as comprehensive. Nevertheless, this
essay has demonstrated that many suggestions for reform are incomplete as they
do not fully account for individual behaviour within a system and how this can
affect the outcome of substantive rules.

SAM PRICE
**To what extent have the cases of Viking, Laval and Demir led to change in the ability of trades union to take effective strike action in Europe?**

ALEX JUST

**INTRODUCTION**

Despite over sixty years of increasing economic, political and social integration the European Union has no harmonized labour law system. This has created a fundamental conflict between the competing European goals of establishing a globally competitive free market and ensuring that all EU citizens live in nation states that respect the progressive ‘European social model’. In the wake of the 2004 EU enlargement the vexed question of ‘social dumping’ has became a more prominent legal issue. As the European labour market expanded, thousands of posted workers from Eastern Europe moved to more economically prosperous member states creating an array of industrial relations conflicts. It is out of such conflicts that the European Court of Justice and the European Court of Human Rights were tasked with attempting to adjudicate on whether the right to strike was consistent with the principles of economic freedom that have consistently been at the heart of the European legal system.

The extent to which the cases of Viking, Laval and Demir have brought about a change in the ability of trades union to take effective strike action in Europe is pertinent because it poses four intriguing questions. First: should EU member states have the right to determine their own social models and labour law in accordance with the established legal principle of subsidiarity or should they be constrained in their approach to labour relations policy by the economic freedoms guaranteed by EU treaties? Second: what should determine the balance between guaranteeing economic freedoms on the one hand and fundamental rights on the other? Third: should industrial action that has the effect of restricting economic freedoms be justified by trades unions before domestic courts, and ultimately before the European Court of Justice? Fourth: to what

---

extent has the recent European jurisprudence had a chilling effect on national and multi-national trades union in Europe?

In order to answer these questions it will first be necessary to outline the basic European legal framework and in particular the primary sources of law that support both the principle of economic freedom and the right to strike. The Viking, Laval and Demir cases will then be analyzed in detail. Finally, the impact recent European jurisprudence has had on trade unions will be discussed. This requires a dissection of the way in which the courts have attempted to balance economic freedoms on the one hand and fundamental social rights on the other. Particular attention will be paid to whether the recent European case law has had a chilling effect on trade unions in the United Kingdom. The broader role of the courts in the context of European industrial and labour relations will also be reviewed. It will become clear that although there has been an effect on some European trade union’s ability to take effective strike action it is as yet uncertain if there will be a long-term chilling effect due to the ambiguous future role national courts will play in interpreting the ECJ's new Viking-Laval proportionality test.

**PART ONE: Legal Principles**

**(i) The European legal order**
A complex and sophisticated legal framework has evolved in Europe involving a combination of multi-lateral treaties and binding secondary legislation in the form of regulations, directives, opinions and recommendations. In order to provide context for the emerging conflict between economic freedoms and the right to strike we must briefly examine the three most relevant sources of European legal power.

(a) The Treaty on European Union (EC), more commonly referred to as The Lisbon Treaty, and the Treaty on the Functioning of the European Union (TFEU) establish many of the principles around which European institutions function. For our analysis the two most important of these ‘constitutional’ legal doctrines
are the *principle of subsidiarity* and the *Article 267 reference procedure*. The principle of subsidiarity is defined in Article 5 EC. It requires decision-making bodies with responsibility for larger areas to perform only those functions that decision-making bodies with responsibility for smaller areas cannot fulfil themselves. For example, the Treaty requires the Community to take action, "only if and in so far as the objectives of the proposed action cannot be sufficiently achieved by the Member States [and can] by reason of the scale or effects of the proposed action be better achieved at Union level". Therefore, if an issue which affects the realisation of an EU objective can be better addressed through national, rather than European, policies and legislation, then it is incumbent upon the EU to allow national governments to resolve the issue. In terms of industrial relations conflicts in Europe, subsidiarity means that as far as possible legal decisions should be taken at the lowest appropriate level, that is, as close as possible to the people who are affected by those decisions. In terms of its practical implication for labour lawyers, as Kenner notes, "subsidiarity has... reinforced the powers of the Members States at the expense of the Community by creating an assumption that, in areas of shared competence, the appropriate level of action is national."

Article 267 TFEU, established a process for national courts in EU member states to make references on matters of interpretation of EU law to the European Court of Justice (ECJ) based in Luxembourg. It is a co-operative rather than an appellate procedure, meaning that once an unresolved question of EU law arises in a national court, the judges in that court must suspend proceedings and send a series of questions to the ECJ. Once the ECJ has made their interpretation it is for the national court to apply the judgment. This is significant because the ECJ cannot proactively answer questions of interpretation, and this strict policy of non-interference means that the all-controversial labour law issues discussed below have been generated by member states, rather than by any overarching European legal actor.

---


(b) The Charter of Fundamental Rights of the European Union (CFR) is a document containing human rights provisions, ‘solemnly proclaimed’ by the European Parliament, the Council of the European Union, and the European Commission in December 2000. In 2007, following the ratification of updated Treaty on European Union in Lisbon; the Charter became a legally binding document in all EU member states except Poland and the United Kingdom whose government’s secured comprehensive opt-outs. The Charter effectively incorporates the European Convention of Human Rights (ECHR) into the EU legal order and makes judgments of the European Court of Human Rights (ECtHR) in Strasbourg legally binding on EU institutions. However, it is important to note that Article 51 CFR states that the Charter is only directly effective regarding EU institutions and the implementation of EU law, meaning national employment legislation might not allow applicants to trigger the more pro trade union seam of ECHR case-law.

(c) The European Social Charter 1996 (ESC) is a Council of Europe treaty that was adopted in 1961 and revised in 1996. The Charter lays out comprehensive human rights and freedoms, several relating to the industrial relations arena. Moreover, there is a supervisory mechanism similar to that deployed at the International Labour Organisation (ILO) guaranteeing their respect by ratifying member States. State parties to the Charter must submit annual reports showing how they have implemented the provisions of the document successfully in law and practice. The European Committee of Social Rights (ECSR) is the body charged with responsibility for monitoring State compliance.

(ii) Economic freedom in Europe
Historically encouraging free trade and reducing anti-competitive business practices has been at the core of the European legal framework. The principle of

---

4 Not to be confused with the European Council, the Council of Europe is an international organisation that promotes co-operation between all countries of Europe in the areas of legal standards, human rights, democratic development, the rule of law and cultural co-operation. It was founded in 1949, has 47 member states (including Turkey and Russia) and is an entirely separate body from the European Union.
the freedom of establishment has been strongly codified in Articles 49 to 55 of the Treaty on the Functioning of the European Union. Article 49 TFEU states:

"restrictions on the freedom of establishment of nationals of a Member State in the territory of another Member State shall be prohibited... Freedom of establishment shall include the right to take up and pursue activities as self-employed persons and to set up and manage undertakings, in particular companies or firms within the meaning of the second paragraph of Article 54."

The right of establishment is granted both to natural and legal persons by the second paragraph of Article 54 TFEU, which defines businesses as, “companies or firms constituted under civil or commercial law, including cooperative societies, and other legal persons governed by public or private law, save for those which are non-profit-making.” Furthermore, Article 56 TFEU guarantees that no business should have restrictions placed on their ability to provide services in the Member States of the EU.

(iii) The right to strike in Europe

Article 153 (5) of the Treaty on the Functioning of the European Union appears to explicitly exclude the right to strike from the EU’s competence stating that, “The provisions of [interventionist social policies by the EU] shall not apply to pay, the right of association, the right to strike or the right to impose lock-outs.”

However, the right of trade unions to take collective action can in fact be found in various European legal sources. First, Article 28 of the Charter of Fundamental Rights of the European Union on the rights of collecting bargaining and action states that,

"Workers and employers, or their respective organisations, have, in accordance with Community law and national laws and practices, the right to negotiate and conclude collective agreements at the appropriate levels and, in cases of conflicts of interest, to take collective action to defend their interests, including strike action.”

---

5 Article 49, Treaty on the Functioning of the European Union at http://euwiki.org/TFEU
6 Article 56, Treaty on the Functioning of the European Union at http://euwiki.org/TFEU
Second, the right is guaranteed by Article 6(4) of the European Social Charter, which declares,

"With a view to ensuring the effective exercise of the right to bargain collectively, the Parties undertake: the rights of workers and employers to collective action in cases of conflicts of interest, including the right to strike, subject to obligations that might arise out of collective agreements previously entered into."\(^9\) (emphasis added)

The right can be deemed fundamental in the sense that, without providing for an absolute guarantee, the Charter only allows for narrow restrictions. Pursuant to the Appendix to the Charter, Article 6(4) ESC can be regulated only once several preconditions have been met and any restriction must be prescribed by law, pursue a legitimate purpose and be necessary in a democratic society. Third, Article 11 of the European Convention of Human Rights outlines various rights of association including that,

"1. Everyone has the right to freedom of peaceful assembly and to freedom of association with others, including the right to form and to join trade unions for the protection of his interests.

2. No restrictions shall be placed on the exercise of these rights other than such as are prescribed by law and are necessary in a democratic society in the interests of national security or public safety, 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. This article shall not prevent the imposition of lawful restrictions on the exercise of these rights by members of the armed forces, of the police or of the administration of the State."\(^10\) (emphasis added)

As will become clear, the case law of the European Court of Human Rights links the ability of trade unions to protect the interests of their members with a collective bargaining framework that includes a right to strike. It is notably that in paragraph 2 of Article 11 ECHR it is permissible for states to restrict the collective action rights of certain groups of employees, primarily those who are deemed necessary to guarantee national security.


\(^10\) Article 11, European Convention of Human Rights at http://www.hri.org/docs/ECHR50.html
PART TWO: European Jurisprudence

(i) International Transport Workers Federation v Viking Line ABP

Facts

A Finnish company Viking Line ABP ran a ferry called The Rosella between Finland and Estonia. Viking wanted to operate the ship under an Estonian flag so it could employ Estonian workers on lower wages than the higher Finnish wages it paid the current crew. The International Transport Workers Federation (ITWF) had a policy to oppose such “reflagging” for the convenience of firms registering their vessels abroad in a low-labour cost jurisdiction, when their real base of operations was in another state. As a consequence of Viking’s actions the Finish Seamen’s Union (which was an ITWF member) planned strike action. Moreover, the ITWF informed its affiliate unions across Europe not to enter into any negotiations with Viking and to hinder its business where possible. Viking consequently sought injunctive relief in the English High Court against the ITWF (gaining standing because the ITWF’s headquarters were in London) on the grounds that the proposed industrial action would directly infringe its right to freedom of establishment under Article 49 TFEU. At first instance the injunction was granted and the strike could not go ahead. However, the ITWF appealed to the English Court of Appeal who overturned the injunction (on the basis of a long-held English legal test on the balance of convenience\(^\text{11}\)) and made an Article 267 reference to the European Court of Justice on the basis that there was an important issue of EU law to be decided which affected, “the fundamental rights of workers to take industrial action.”\(^\text{12}\)

Decision

The ECJ held that although it was for every national court to ultimately adjudicate on the legality of strike action it was possible that collective action such as that undertaken by the ITWF could be unlawful under European law

\(^{11}\) See: American Cyanamid Co (No 1) v Ethicon Ltd [1975] UKHL 1

because it infringed the employer’s rights to freedom of association under Article 49 TFEU. The ECJ acknowledged that, “the right to take collective action, including the right to strike, must… be recognised as a fundamental right which forms an integral part of the general principles of Community law.” 13 However, the court also stated that, “the exercise of [the right to strike] may none the less be subject to certain restrictions… in accordance with Community law and national law and practices.” 14 Moreover, in order for a strike to justifiably infringing the economic freedoms of another party then the workers for whom a trade union was acting must have their jobs or conditions of employment “jeopardised or under serious threat”.15 In summarising its judgement the ECJ laid a proportionality test, which declared that it was the responsibility of national courts to confirm that any strike action “does not go beyond what is necessary.”16

Analysis
This judgment led many to conclude that the ECJ judges had taken a “double approach” in that they adopted the principle of striking as a fundamental right in concert with an internal-market balancing act. Hendrickx has even gone so far as to argue that, “labour law could be seen as a justified limitation to European internal market law”17 rather than as a positive body of law that lays out the rights of all worker and trade unions in the EU. In the influential Albany judgement18 the ECJ had shown itself to be willing to exempt a trade union from the free movement treaty provisions in the interests of achieving a broader social goal.19 The ECJ’s resolute unwillingness to extend an Albany like exception to strike action was understandably troubling for the trade union movement. However, shortly after the Viking judgement a case involving similar legal

14 Ibid. Para 81.
15 Ibid. Para 75.
16 Ibid. Para 75.
19 In Albany the social good was deemed to be the provision of a group pension scheme for Dutch textile workers that clearly fell foul of EU anti-trust regulations.
arguments reached the ECJ that many hoped would see the jurisprudential pendulum swing back in favour of striking workers.

**(ii) Laval Un Partneri Ltd v Svenska Byggnadsarbetareförbundet**

**Facts**

A Latvian construction firm, Laval Un Partneri Limited were awarded a contract by the Swedish government to renovate a number of high school premises. Laval sent Latvian workers to work on these sites (which were operated by a Swedish company that Laval itself owned) where they received considerably lower wages compared to their Swedish counterparts. The Swedish Building Workers’ Union asked Laval to sign its collective agreement that covered all workers on state construction projects. This collective agreement would have resulted in more favourable conditions for the Latvian “posted workers” than those expected by European law under the *Posted Workers Directive* and also contained a clause, which would prevent Laval from determining pay rates in advance of establishing operations in Sweden. Management at Laval refused to sign the collective agreement and in response the Swedish Building Workers’ Union (supported by the largest Swedish electricians union) called a strike to blockade Laval’s building sites. Consequently, Laval was unable to honour its contractual obligations in Sweden and the firm’s subsidiary went bust. In an attempt to seek compensatory damages and a declaration that the strike was unlawful, Laval argued that the action of the Swedish Building Workers’ Union violated Article 56 TFEU, which prohibits restrictions on freedom to provide services between member states. As in *Viking* the domestic Swedish court sent the matter to the ECJ using the Article 267 referral mechanism.

**Judgement**

Following on from the logic of its judgement in *Viking*, the ECJ reiterated the fact that trade unions do have a, “right to take collective action for the protection of the workers of the host state against possible social dumping may constitute an

---

overriding reason of public interest”\textsuperscript{21} and that this right could potentially justify an infringement of an employer’s economic rights. However, in \textit{Laval}, the ECJ ruled that a private undertaking could invoke Article 56 TFEU against a trade union, since their freedom to provide services would be compromised if the abolition of State barriers could be neutralised by obstacles resulting from the legal autonomy of associations or organisations. This led the Court to conclude that the union’s blockade could not be justified on the grounds of protecting workers. The ECJ was also concerned about the unpredictability that the collective agreement placed on Laval, in so far as wage negotiations had to be thrashed out on a local level once the company was already operating in Sweden. In general the ECJ was critical of the Sweden’s collective bargaining framework and clearly felt it to be too imprecise to allow any firm to quantify its financial obligations in advance,

“collective action such as that at issue in the main proceedings cannot be justified in the light of the public interest objective... where the negotiations on pay, which that action seeks to require an undertaking established in another Member State to enter into, form part of a national context characterised by a lack of provisions, of any kind, which are sufficiently precise and accessible that they do not render it impossible or excessively difficult in practice for such an undertaking to determine the obligations with which it is required to comply as regards minimum pay.”\textsuperscript{22}

Furthermore, the ECJ developed the \textit{Viking} proportionality test by adding that strike action had to be suitable for attaining the ‘legitimate public interest’ objective it pursued as well as not going beyond what was necessary to achieve this goal. The net effect of these changes in the industrial action context is that the more a strike restricts the employer’s free movement rights, and therefore

\textsuperscript{21} C-341/05 \textit{Laval un Partneri Ltd v Svenska Byggnadsarbetareförbundet, Svenska Byggnadsarbetareförbundets avd. 1, Byggetan, Svenska Elektrikerförbundet [18/12/2007] Para 103.

\textsuperscript{22} C-341/05 \textit{Laval un Partneri Ltd v Svenska Byggnadsarbetareförbundet, Svenska Byggnadsarbetareförbundets avd. 1, Byggetan, Svenska Elektrikerförbundet [18/12/2007] Para 110.
the more effective it is from a trade union perspective, the harder it will be to justify before the courts.23

Analysis
Whereas in Viking the ECJ appeared to leave the final decision on proportionality to the national court in each member state, on the facts of this case the court was willing to declare that the Swedish union had acted disproportionately and in such a way that unlawfully restricted Laval’s freedom to provide services. The ECJ’s logic in Laval would seem to demand that trade unions consider additional measures to avoid taking collective action beyond those contemplated by statute. Novitz speculates that this may lead national courts to demand that trade unions make a reference to a consolation or arbitration panel, or provide periods of notice over and above the statutory procedural minima before initiating a strike.24

It must also be noted that the Laval judgement is complicated by the factual and legal matrix out of which it arises - notable in respect of the Posted Workers Directive. According to Hendrickx, the purpose of the Posted Workers Directive was to provide a mechanism, “to protect workers against social dumping, to ensure fair competition between companies, and to respect certain Member State’s (policy choices regarding) minimum provisions.”25 It is a fundamental principle of European Union law that Directives cannot impose obligations on private parties (i.e. Directives do not have horizontal direct effect) but in the Laval judgement the ECJ stated that they had used the Posted Workers Directive as a guide to the interpretation of broader Treaty articles. Davies has even gone so far as to suggest that, “the ECJ’s real concern in Laval is with the way in which the Swedish government had chosen to implement the Posted Workers Directive”

and that we should therefore be hesitant about drawing wider conclusions from the case. However, it is clear from the judgment that the ECJ found that free movement rights regarding establishment and services can have horizontal application, which means they could be effective against trade unions.

At this stage, it is worth highlighting the two most significant differences between the Viking and Laval decisions. First, in Viking it was accepted that preserving jobs and employment conditions was a legitimate objectives of strike action but only if workers were under ‘serious threat’. This appeared to allow for ‘defensive’ collective action by trade unions without permitting strikes for improving existing terms and conditions of work. However, in Laval the latitude of legitimate strike action was broadened by the ECJ. The scope of a trade union’s objectives could be to ensure the “protection of workers” in more general terms, although the court stopped short of legitimizing secondary solidarity strikes or pro-active politically minded collective action. However, as had been noted, this apparent dissimilarity is less acute if we read the Laval judgment solely in the context of the Posted Workers Directive. Second, in Viking the ECJ appeared to leave it to the discretion of national courts to determine whether collective action was proportionate, whereas in Laval it decided to turn its hand to analyzing the facts of the case in order that it could outline a more sophisticated model for assessing the legitimacy of the Swedish Builder’s Union strike. As Davies notes, “the rulings in Viking and Laval have the potential to involve the courts in a much more politically sensitive set of questions.” 27 It will be fascinating to observe in the coming years how both the ECJ and national courts seek to apply the proportionality test to unions’ strike action.

(iii) Demir and Baykara v Turkey

Facts

Ms Demir and Mr Baykara were both members of Tüm Bel Sen, the Turkish trade union for civil service employees. In 1993 the union signed a two-year collective

---

26 davies, a.c.l. (2008), ‘one step forward, two steps back? the viking and laval cases in the ecj’, 37 in industrial law journal, p.137.
bargaining deal with all Turkish local authorities. However, the Gaziantep Municipal Council failed to comply with the pay-rise escalator they had agreed with Tüm Bel Sen and Demir and Baykara brought proceedings in the District Court alleging breach of contract. Having won their case at first instance, the local authority appealed to the Court of Cassation, which quashed the decision holding that although Demir and Baykara had a right to join a union, Tüm Bel Sen had no authority to enter into collective agreements under Turkish law. The case was remitted back to the District Court, which defiantly overruled the Court of Cassation and restated its view that Turkish civil servants did have a right to enter into collective agreements through their registered trade union on the basis that this line of legal logic was consistent with Turkey’s ratification of various ILO conventions and certain sections of the European Social Charter. Unfortunately for the applicants, their case bounced back to the Court of Cassation for a second time, where the District Court’s decision was again overturned. To make matters worse for Demir and Baykara, a separate claim was brought against Tüm Bel Sen in the Turkish Audit Court, who found that not only did civil servants have no authority to participate in any form of collective bargaining but that all their members had to repay the additional benefits they had received from the Gaziantep Municipal Council during the period that the (now defunct) collective agreement was in operation. In 1996, after all their domestic appeals were exhausted, the trade union applied to the European Court of Human Rights alleging a breach of freedom of association under Article 11 ECHR and protection against discrimination under Article 14 ECHR.

The Strasbourg court does not have a reputation for dispensing speedy justice, so it was not until 2006 that a Chamber of seven judges heard the case of Demir and Baykara v Turkey. The ECtHR initially held that Turkey had no case to answer in respect of the Article 14 ECHR discrimination claim, but that the Turkish Government was guilty of breaching Article 11 ECHR by denying the applicants access to collective bargaining rights. Turkey appealed against the initial Chamber’s ruling and the case was referred to the final appellate body of the ECtHR – the Grand Chamber.
Judgment

The Grand Chamber held unanimously that the actions of the Gaziantep Municipal Council had been a disproportionate and unjustified interference with the applicant’s right to freedom of association. Although it acknowledged that it was permissible for the Turkish government to place restrictions on the trade union activities of certain groups of employees (as per Article 11 (2) ECHR), even under Turkey’s restrictive civil service personnel codes there was no provision that justified an out-right ban on union activity. Moreover, the Grand Chamber gave more general advice on the nature and application of the ECHR exception stating that,

"it must also be borne in mind that the exceptions set out in Article 11 are to be construed strictly; only convincing and compelling reasons can justify restrictions on such parties’ freedom of association. In determining in such cases whether a "necessity" – and therefore a "pressing social need" – within the meaning of Article 11 (2) exists, States have only a limited margin of appreciation, which goes hand in hand with rigorous European supervision embracing both the law and the decisions applying it, including those given by independent courts.”28

The court then assessed whether or not the Court of Cassation had been right to (twice) annul the collective agreement negotiated by Tüm Bel Sen and the local authority. To the surprise of many, the Grand Chamber answered this question by outlining what it deemed to be the constituent elements of an effective labour relations system that complied with the European Convention on Human Rights. The court asserted that the Article 11 guarantee of freedom of association created a multi-faceted right, which had to incorporate means for trade unions to engage in collective bargaining on behalf of its members. Moreover, any national collective bargaining framework was said to be functionally useless if trade unions did not have the ability to use strike action to gain leverage at the negotiating table. The Grand Chamber also ruled that the ECHR should be viewed in the context of evolving European social and legal norms.

28 Demir and Baykara v Turkey (2009) 48 EHRR 54, para 19.
Analysis

By adopting a ‘living document’ model of construction in respect of the ECHR, the right to strike now has to be interpreted by referencing international labour standards (notable ILO Conventions 98 and 151\(^{29}\)), Article 6 (4) of the European Social Charter, Article 28 of the Charter of Fundamental Rights of the European Union and “the practice of European States”.\(^{30}\) This has the unprecedented effect of making ILO conventions binding on European states even in cases where national governments had chosen not to ratify particular ILO conventions. The Grand Chamber justified this decision on the basis that,

“...Being made up of a set of rules and principles that are accepted by the vast majority of States, the common international or domestic standards of European States reflect a reality that the Court cannot disregard... in searching for common ground among the norms of international law it has never distinguished between sources of law according to whether or not they have been signed or ratified by the Respondent State.”\(^{31}\)

A corollary of the Demir decision was that rulings of the ILO Committee of Experts were now deemed to be persuasive in European employment law cases. Ewing and Hendy highlight the significance of this development arguing that,

"In treating the ECHR as a living instrument, the Strasbourg court is also acknowledging that these other treaties are living instruments as well, in the sense that in considering their scope and content, it is necessary to have regard not only to the text of the treaties but also the jurisprudence of their supervisory bodies.”\(^{32}\)

The internal logic of the Demir decision was upheld by the Grand Chamber in the 2009 case of Enerji Yapi-Yol Sen v Turkey\(^{33}\), which further widened the gulf between the ECHR jurisprudence and that of the ECJ following Viking and Laval. This had led many to predict a showdown between the Strasbourg and Luxembourg courts. Ewing and Hendy have argued that,

"It is difficult to see how the European Court of Human Rights could avoid upholding Article 11 and the right to collective bargaining and to strike over the business freedoms contained in what are now Articles 49 and 56 of the TFEU.

---

\(^{29}\) See [http://www.ilo.org/ilolex/english/convdisp1.htm](http://www.ilo.org/ilolex/english/convdisp1.htm)

\(^{30}\) Demir and Baykara v Turkey (2009) 48 EHRR 54, Paras 98-101 and 147-151.

\(^{31}\) Ibid. paras 78 and 85.


\(^{33}\) Enerji Yapi-Yol Sen v. Turkey (Application No. 68959/01).
And so issues would bat to and fro between the two courts in a titanic battle of the juristocrats, each vying for supremacy in the European legal order.”

The broader, and as yet unanswered, ‘constitutional’ question for EU member states following Demir and Baykara and Enerji Yapi-Yol Sen, is whether ECHR case law is going to be more persuasive in national courts. In theory, under the Treaty of Lisbon legal order the EU Charter on Fundamental Rights will allow domestic judges to read down ECHR (and now by association ILO) convention rights when adjudicating on employment law disputes at first instance. However, as will become clear, in practice the reaction of European courts has been decidedly mixed on this issue and as yet the right to strike is far from being considered an unbending legal principle.

**PART 3: Impact on Trade Unions**

(i) A Balancing Act?

The analysis of Viking, Laval and Demir demonstrates that thus far the European courts have been unable to definitively explain what should determine the balance between guaranteeing economic freedoms on the one hand and fundamental social rights on the other. It seems from the current case law that the balance between both types of rights will change according to the respective relative weight placed on them by the various sources of European legal authority. In political terms, the Treaty of Lisbon made it clear that the EU wishes to realize a “Social Europe”. For this reason labour lawyers should be able to deploy arguments that rely on a number of broad international socio-legal instruments – notable the European Convention on Human Rights and the European Social Charter. Moreover, following Demir ILO Conventions are also persuasive, even against non-ratifying governments, which would tend to suggest that the balance is aligned in favor of a pro-union stance that recognizes the right to strike as the bedrock of any legal labour-relations framework.

However, the ECJ framed their decisions in *Viking* and *Laval* in such a way that they placed a far greater emphasis on the economic freedoms under EU law and analysed collective action not as a positive social right, but rather as an occasionally justifiable mechanism to infringe on a firm’s sacrosanct free movement rights. As Barnard rightly notes, “the very fact that the collective action is found to be a restriction... and thus in principle a breach of Community law automatically puts the ‘social’ on the back-foot.”\(^{35}\) Although the Court recognized the right to strike as fundamental, it did so in what Davies terms a “defensive context”.\(^{36}\) Hendrickx has supported this logic by arguing that the recent judgements, “are as much about European ‘internal market law’ as about trade union or strike law.”\(^{37}\) The question of whether these rhetorical nuances illuminate a substantive and consistent jurisprudence on the part of the European courts was in part answered by the 2010 case *Commission v Germany.*\(^{38}\)

One of the first accounts of how the balancing act should be performed in light of *Viking, Laval* and *Demir* is found in the Advisory Opinion of Advocate General Trstenjak in *Commission v Germany*. AG Trstenjak acknowledged the need for European courts to adopt a symmetrical approach to such reconciliation. Therefore, in cases where there is a conflict between a social right and an economic freedom both legal positions must be presumed to have equal status due to the enhanced importance of social rights in the Lisbon Treaty and the Charter of Fundamental Rights of the European Union. The Opinion also stressed the fact that just as fundamental freedoms can be used to justify restrictions on a fundamental right, it is equally possible that a fundamental right can lead to the limitation of the scope of a fundamental freedom.\(^{39}\) However, *Commission v...
Germany was silent on the issue of whether the new Viking-Laval proportionality test now had to be read in light of ILO Conventions and the opinions of the ILO Committee of Experts. Nicol observes that while we cannot predict with any confidence the fate of the Viking-Laval case one option for the ECJ may be to push back against the ECHR seam of cases and modify its jurisprudence to create a “greater respect for internal state value spaces.” This may grant member states more flexibility and autonomy in deciding on the legitimacy of strike action, as was presumably envisaged by the framers of the current EU legal order given the clear language of Article 153 (5) of the Treaty on the Functioning of the European Union which excludes the regulation of collective action by EU institutions.

ii) A Chilling Effect?
Recent developments in the United Kingdom serve as a helpful case study to assess the extent to which the recent European jurisprudence has had a chilling effect on national and multi-national trades unions. First, the Viking-Laval proportionality test has not been uniformly applied at first instance by British judges. This is due in part to the way in which the European case law interacts with the domestic interim injunctions application that is filed by an employer hoping to avert strike action that has been legally called by a trade union. The classic American Cyanamid test requires an employer to first prove that there is a serious issue to be tried. This hurdle is generally overcome, as an employer is able to demonstrate to the court that were the proposed strike action to proceed it would have a significant negative financial impact on their business. The second key question to address is who does the balance of convenience favour? In deciding which party will suffer the greater harm if the injunction is or is not granted, British courts are not required to undertake any

---

41 For more on the effective of recent legal develops in other Member States see the country specific essays on Belgium, France, Germany, Italy, Spain and the Netherlands in Ales, E. And Novitz, T. (eds.) (2010) Collective Action and Fundamental Freedoms in Europe. Oxford, UK.
42 American Cyanamid Co (No 1) v Ethicon Ltd [1975] UKHL 1 (05 February 1975).
analysis of the merits of the strike itself. This means that the UK courts can effectively ignore the recent ECJ cases and in particular the *Viking-Laval* proportionality test. Breach of such an injunction (even if it was granted in a curious British legal vacuum that ignores European precedent) is treated as contempt of court under UK law, which means that trade unions that decide to proceed with a strike can face severe criminal penalties.\(^{43}\)

Moreover, it would seem that recent European case law could result in the removal of the statutory damages cap that has been the cornerstone on UK trade union law since the 1970s. Under the *Trade Union and Labour Relations (Consolidation) Act 1992 (TULRC)* trade unions that are found to have conducted an illegal strike are only liable in tortuous damages up to a fixed level depending on their level of membership. Apps argues that under EU law it may be possible (as with sex discrimination cases) for employers to argue that they are being denied an effective remedy if there is a capped upper award in a civil tort claim.\(^{44}\)

There is evidence from a recent dispute involving the British Air Line Pilots Association (BALPA) that these developments have indeed had a chilling effect. Novitz asserts that a planned BALPA strike in 2008, shortly after the publication of the *Viking* and *Laval* judgements, was called off due to the union fears that a UK court might be unsympathetic to their attempt to limit the economic freedoms of various European airlines and the crippling inflated damages claims that may follow as a result of breaching a High Court injunction.\(^{45}\) It has also significant that under TULRC secondary strikes allowed in the UK are banned which may present problems in the future for international or European level collective agreements and UK trade unions might not be able to stand in solidarity with their European brothers and sisters.

---


(iii) The future role of courts?

The UK experience illuminates the nuanced, and at times fraught relationship between national courts and the European Court of Justice in the employment law arena. As O’Learly notes, “although it is in the [ECJ’s] interest to preserve the semblance of an equal and co-operative relationship with national courts, it is clear that, as regards the interpretation and validity of EC law, it is in a hierarchically superior position.” If the European legal order now guarantees protection of social rights (including the right to strike) through its jurisprudence, Directives or by allowing trade unions to make ILO Conventions binding on national government, then it seems logical that it should be mandatory for a court to balance the economic freedoms enshrined in the EC Treaty against such social rights. Applying the principle of subsidiarity, it follows that this process should be undertaken at first instance by national courts.

Although this is evidently not happening at present in the UK, it is important to consider whether the Viking-Laval proportionality test will grant judges in Europe too much discretion. It is arguable that the greater emphasis on this legal test may end up with a concrete legal right to strike running into difficulties. As has been discussed the, second limb of the proportionality test in Laval states that trade unions can only take strike action up to the point that is necessary for them to achieve their final ‘public interest’ objective. There is a serious danger that either national or European courts will attempt to identify alternative dispute resolution paths in an industrial conflict without considering their effectiveness in a specific bargaining process. This more interventionist role for courts in Europe may well prove highly restrictive of the right to strike.

---

CONCLUSION
The line of cases discussed has confirmed that trade unions in the European Union have a fundamental legal right to take collective action. Davies is, therefore, right to observe that, “the ECJ’s position may be of some value on a policy level and has accordingly been welcomed by the [UK] Trades Unions Congress.”48 However, it is ambiguous under which conditions collective action is ‘legitimate’ in the broader context of an EU legal framework that places so much emphasis on the economic freedoms of firms. The biggest problem for the European trade union movement is that Viking and Laval only allow ‘defensive’ strikes to take place in order to protect the immediate position of workers whose jobs are at risk or serve a ‘public interest objective’. This effectively rules out all secondary collective action or political strikes in solidarity with other actors in the European labour movement. Moreover, this problem is compounded from the perspective of trade unions by the fact that the more effective their industrial action is likely to be the more damage it is likely to cause the employer and thus the more likely it is for courts to feel their strike is disproportionate and illegal under EU law.

Furthermore, the Viking-Laval proportionality test places a great deal of pressure on national courts to decide if trade unions have ‘struck the right balance’ and whether they could have achieved their ‘legitimate’ aims without resorting to full-scale industrial action. In EU member states like the UK and Ireland national courts have tended to take a hands-off attitude to adjudicating on the legitimacy of strike action, preferring a more technical statute-based approach. However, in Germany and Belgium judges use proportionality tests frequently. This means that the impact of the Viking and Laval decisions is going to be felt differently across the EU.

In terms of a chilling effect we can reach three firm conclusions. First, it is still too early to tell if all of the ECJ’s judgement in Laval is going to be persuasive in national courts outside the narrow context of the Posted Workers Directive.

Second, it is equally unclear if Demir and the subsequent Grand Chamber jurisprudence is going to be binding in a new Lisbon legal order in which the ECHR and the EU Charter on Fundamental Rights is more authoritative. Moreover, we will have to wait to see if EU member states really start paying serious attention to ILO Conventions and the findings of the Committee of Experts who have in the past consistently chastised states for falling short in ensuring basic trade union freedoms. Third, it is still essential to consider how each national labour relations framework impacts the ability of trade unions to take effective strike action. For instance, in the UK, the likely removal of the statutory cap on damages imposed on trade unions in cases of ‘illegal’ strikes ultimately did deter BALPA workers from striking against their employer.

Overall, Viking, Laval and Demir have had a major impact on the risk calculus for trade unions in Europe but we must wait to observe if and how the ECJ’s new proportionality test is applied in national courts to establish whether there will be a genuine long-term chilling effect.

ALEX JUST
Intellectual Property Law & the needs of consumers, creators / innovators and the public interest: An appropriate balance?

ANDREW PELLINGTON

There have been numerous arguments based on the extent to which Intellectual Property law balances the needs of consumers, creators and the public interest at large. To explore these debates, it is necessary to investigate the purpose of intellectual property law, along with how the law tries to balance the needs of consumers, creators/innovators and the public interest. The extent to which this is adequately achieved will also be considered.

What is Intellectual Property Law?

The Law of Intellectual Property acts as a form of protection for the results produced from human creativity.1 It therefore, gives creators exclusive rights to exploit their creation,2 whilst discouraging others from taking unfair advantage of the creation.3

This protection can be justified in the context of the Labour theory derived from the 17th Century philosopher John Locke who acknowledged that the basis of protection provides the creator with a natural right to control the use of their creation,4 whilst simultaneously rewarding the creator for the effort that has been put into the creation through protection, thereby satisfying the maxim 'he who sows shall reap'.5

---

2 Hegel G., W., F., Philosophy of Right (Ontario: Batoche Books Ltd, 2001) p 61, para. 50: 'It is a self-evident...that an object belongs to him who is...first in possession of it. A second person cannot take into possession what is already the property of another.'
3 World Trade Organisation, 'Intellectual property: protection and enforcement: Origins: into the rule-based trade system' at http://www.wto.org/english/theWTO_e/whatis_e/tif_e/agrm7_e.htm accessed on 14th April 2010: 'Creators can be given the right to prevent others from using their inventions, designs or other creations — and to use that right to negotiate payment in return for others using them. These are "intellectual property rights"...'
4 Locke, J., Two Treatises of Government (Cambridge: Cambridge University Press, 1988) p 287-288, para. 27: '...every Man has a Property in his own Person. This no Body has any Right to but himself. The Labour of his Body, and the Work of his Hands...are properly his...'
5 Victoria Park Racing Co v. Taylor (1937) 58 CLR 479 at 509, per Dixon J: ‘...[The courts] have not in British jurisdictions thrown the protection of an injunction around all the intangible elements.
However, such protection requires a balance to be made between the interests of consumers, the creator, and of the public at large.\textsuperscript{6} Indeed, in \textit{Chiron v. Murex Diagnostics},\textsuperscript{7} Aldous J observed that a monopoly right granted to an inventor, provides them with a right to restrict competition and place higher prices for the use of their invention which will ultimately be contrary to the public interest.\textsuperscript{8}

Consequently, a balance is required in the law, between maintaining consumer interests in promoting fair competition,\textsuperscript{9} along with the general public interest in acquiring a right of access to use the creators work.\textsuperscript{10} Indeed, the need for balance is also identified at an international level, under Art. 27(2) Universal Declaration of Human Rights,\textsuperscript{11} which provides a public right of enjoyment from a creators work.\textsuperscript{12}

\textbf{Consumers}

The law of trademarks allows consumers to distinguish between one traders goods or services from those of another, as a result of the registered mark indicating the origin of goods or services.\textsuperscript{13} The trade mark is therefore required

\begin{flushleft}
\textit{of value...in exchange, which may flow from the exercise by an individual of his powers or resources whether in the organisation of a business or undertaking or the use of ingenuity, knowledge, skill or labour. This is sufficiently evidenced by the history of the law of copyright and by the fact that the exclusive right to invention, trade marks, designs, trade name and reputation are dealt with in English law as special heads of protected interests and not under a wide generalisation...'}
\par
\par
\textsuperscript{7} Chiron Corp v. Murex Diagnostics Ltd (No. 10) [1995] FSR 325.
\par
\textsuperscript{8} Ibid. at 332.
\par
\par
\par
\textsuperscript{11} Universal Declaration of Human Rights 1948 Art. 27(2): ‘Everyone has the right freely to participate in the cultural life of the community, to enjoy the arts and to share in scientific advancement and its benefits.’
\par
\textsuperscript{12} Id.; \textit{Trade Related Aspects of Intellectual Property Rights Agreement (TRIPS) 1994 Art. 4(1): ‘With regard to the protection of intellectual property, any advantage, favour, privilege or immunity granted to a Member to the nationals of any other country shall be accorded immediately and unconditionally to the nationals of all other Members...’}
\par
\textsuperscript{13} Trade Marks Act 1994 s. 1(1); (Case 39/97) Canon Kabushiki Kaisha v. Metro Goldwyn Mayer Inc [1998] ECR 1-5507, para. 28: ‘...the essential function of the trademark is to guarantee the identity of the origin of the marked product to the consumer or end user by enabling him, without any possibility of confusion, to distinguish the product or service from others which have another origin. For the trade mark to be able to fulfil its essential role...it must offer a guarantee that all the goods or services bearing it have originated under the control of a single undertaking
\end{flushleft}
to be distinctive otherwise it will be void for registration. Lord Nicholls acknowledged that the distinctive function of the mark is essential to its operation. Indeed, this requirement is also recognised within European legislation through Council Regulation (EC) 207/2009.

The requirement is designed to protect traders as well consumers who are reasonably observant and circumspect, especially as consumers have the power to associate goods or services based on their perception of quality, thereby enabling the trade mark to act as a form of brand identity. Indeed, Rodgers observed that the public perception of brand names play a vital role as to the success of a trader.

...the distinctive character of a sign consisting in the shape of a product, even that acquired by the use made of it, must be assessed in the light of the presumed expectations of an average consumer of the category of goods or services in question, who is reasonably well informed and reasonably observant and circumspect...'; (Case 210/96) Gut Springenheide GmbH v. Oberkreisdirektor des Kreises Steinfurt - Amt fur Lebensmitteluberwachung [1998] ECR I-4657 at para. 31; (Case 53/01) Linde AG v. Deutsches Patent- und Markenamt [2005] 2 CMLR 44 at para. 41.

Scandecor Developments AB v. Scandecor Marketing AB and Others [2001] UKHL 21, para. 18, per Lord Nicholls: '...The ability to apply a mark to goods encourages makers of goods to set and maintain quality standards. It enables customers to make an informed choice between different goods available in the market.'; Re Coca-Cola's Trademark Application; Coca-Cola Ltd v. Pike (No. 2472042) [2010] WL 666343 para. 18; (Case 206/01) Arsenal Football Club Plc v. Reed [2003] Ch 454 at 465, paras 46-47.


In *Arsenal Football Club Plc v. Reed*, it was identified that there may be cases where consumers may be more interested in the mark rather than the goods that the mark is applied to, thus, illustrating the psychological impact of trade marks on consumers. As a result, traders have an interest in protecting the reputation of their mark in terms of maintaining the good name associated with the perceived standard of quality of the good or service offered to consumers, whilst allowing consumers to make informed choices during purchases.

This additional function of trade marks also works simultaneously with other forms of intellectual property rights as it provides trade mark owners with a monopoly right in the mark, in order to prevent unauthorised use of the mark being applied to identical or similar goods or services provided by another trader. Consequently, trade mark protection in a particular product will not restrict competition, and so competitors are able to market their own goods or services, regardless of whether they are identical to the traders, and may lead to consumer confusion.

Although in such cases, it is common for a trader to bring an action in the law of Passing off, which exists to maintain the reputation and good name that has been

---

22 (Case 206/01) *Arsenal Football Club Plc v. Reed* [2003] Ch 454.
23 Ibid. at 466, para. 48.
24 Griffiths, A. 'The impact of the global appreciation approach on the boundaries of trade mark protection' (2001) (4) *Intellectual Property Quarterly* 326-360 at 329: 'Trade marks can gain a "psychological hold" on the minds of consumers, which gives them a selling power above and beyond that of the underlying goodwill. Trade marks which have gained such a hold on the minds of consumers can add significant value to the products with which they are used.'
25 Scandecor Developments AB v. Scandecor Marketing AB and Others [2001] UKHL 21, para. 19, per Lord Nicholls: '...the proprietor of a trade mark [has] an economic interest in maintaining the value of his mark. It is normally contrary to a proprietor's self-interest to allow the quality of the goods sold under his banner to decline.'; *Aristoc Ltd v. Rysta Ltd* [1945] AC 68 at 102; *Glaxo Group Ltd v. Dowelhurst Ltd* [2000] 2 CMLR 571, paras. 17-18.
27 *Trade Marks Act 1994 s. 10(2):* 'A person infringes a registered trade mark if he uses in the course of trade a sign where because—(a) the sign is identical with the trade mark and is used in relation to goods or services similar to those for which the trade mark is registered, or (b) the sign is similar to the trade mark and is used in relation to goods or services identical with or similar to those for which the trade mark is registered, there exists a likelihood of confusion on the part of the public, which includes the likelihood of association with the trade mark.'
developed through the use of the traders mark, and strives to protect the goodwill that has added value to the business through the use of the mark.

Passing off also indirectly protects the interests of consumers from purchasing low quality goods or services in the erroneous belief that they derive from or are provided by another trader. Indeed, in Commissioners of Inland Revenue Appellants v. Muller, Lord Macnaghten claimed that the value of goodwill is worth nothing without the attractive force of consumers.

The practice of comparative advertising, also helps to balance consumer choice as well as allowing traders to differentiate their mark from those of competitors, provided that it is in accordance with honest commercial or industrial practices under s. 10(6) Trade Marks Act 1994.

---

29 Commissioners of Inland Revenue Appellants v. Muller [1901] AC 217 at 223 per Lord Macnaghten at 223-224: ‘...goodwill...is the benefit and advantage of the good name, reputation, and connection of a business. It is the attractive force which brings in custom. It is the one thing which distinguishes an old-established business from a new business at its first start. The goodwill of a business must emanate from a particular...source...’; Draper v. Trist [1939] 3 All ER 513 at 526.

30 Ibid. at 235, per Lord Lindley: ‘Goodwill regarded as property has no meaning except in connection with some trade, business, or calling...the word...[includes] whatever adds value to a business by reason of situation, name and reputation, connection, introduction to old customers, and agreed absence from competition, or any of these things... In this wide sense, goodwill is inseparable from the business to which its adds value, and...exists where the business is carried on.’

31 Commissioners of Inland Revenue Appellants v. Muller [1901] AC 217 at 223.

32 Hodgkinson & Corby Ltd v. Wards Mobility Services Ltd [1994] 1 WLR 1564 at 1570 per Jacob J: ‘At the heart of passing off lies deception or its likelihood, deception of the ultimate consumer in particular...’; Perry v. Truefitt (1842) 49 ER 749.

33 Commissioners of Inland Revenue Appellants v. Muller [1901] AC 217.

34 Ibid. at 223.


37 Comparative Advertising Directive 2006/114/EC Art. 2 (c): ‘comparative advertising means any advertising which explicitly or by implication identifies a competitor or goods or services offered by a competitor’, (Case 112/99) Toshiba Europe v. Katun Germany [2002] 3 CMLR 7, paras. 29-31; (Case 381/05) De Landtsheer Emmanuel SA v Comité Interprofessionnel du Vin de Champagne [2007] Bus LR 1484 at para. 28: ‘Whether undertakings are competing undertakings depends...on the substitutable nature of the goods or services that they offer on the market.’

38 Trade Marks Act 1994 s. 10(6): ‘Nothing in the preceding provisions of this section shall be construed as preventing the use of a registered trade mark by any person for the purpose of
Carty\(^{39}\) highlights that the requirement of trade marks builds trust between consumers as they will associate the product, bearing the mark, to possess a consistent quality, which, in turn, establishes consumer loyalty with the mark.\(^{40}\) She therefore acknowledges that comparative advertising is more concerned with competitive practices rather than providing consumers with information, as the judges tend to view consumer protection through the criminal law via the institutions charged with protecting consumers including the Office of Fair Trading and the Advertising Standards Agency, in cases of misconduct.\(^{41}\) Indeed, s. 20(1) Consumer Protection Act 1987 states that it is an offence for traders to provide consumers with misleading information as to the price of goods or services.\(^{42}\) Although, Carty observes that where a case of serious misinformation applied to a product arises, then the judges may apply trade mark theory to solve the problem.\(^{43}\)

Nevertheless, the Gowers Review of Intellectual Property\(^{44}\) has commented that the existence of comparative advertising, adequately balances the needs of consumer choice and the operation of trade marks in terms of directing consumers towards the traders brand.\(^{45}\) However, problems arise where traders deliberately attempt to take advantage of another traders established trade-mark or goodwill.

'Look-a-like' Packaging
Supermarkets typically sell their own brand name goods alongside other well established brands, thereby increasing the likelihood of consumer confusion.

\(^{40}\) Ibid. at 297.
\(^{41}\) Ibid.
\(^{42}\) Consumer Protection Act 1987 s. 20(1).
\(^{45}\) Ibid. at p 41, para. 3.35.
An example occurred in *United Biscuits (UK) Ltd v. Asda Stores Ltd*, where four pictorial marks depicting penguins had been used for the established Penguin biscuit brand, whilst the defendant had used a similar pictorial and coloured get-up, depicting a Puffin. An injunction was granted to prevent passing off, but the claim for trade mark infringement failed as a result of the differences between the pictorial marks.

In such cases, British law recognises no no from of unfair competition, despite the notion of such competition being prohibited internationally by Member States who are signatories to the Paris Convention 1883. As a result, traders can compete to the extent that they take away another traders market and/or consumers, as well as putting either themselves, or their competitors, out of business through the use of another traders goodwill. Indeed, Lord Scarman in *Cadbury Schweppes v. Pub Squash*, acknowledged that the point at which competition is unfair is unknown, although no wrong is committed where a trader simply enters into a market an competes.

However, Andrew Gowers has observed that where a trader has developed goodwill in a brand name after a 'look-a-like' product has been made, then that

---

47 Id.
48 Ibid. at 526.
49 Ibid. at 513.
50 Swedac Ltd v. Magnet & Southerns Plc [1989] 1 FSR 243 at 249: '...unfair competition is not a description of a wrong known to the law...'
51 Paris Convention for the Protection of Industrial Property 1883 Art. 10bis: '(1) The countries of the Union are bound to assure to nationals of such countries effective protection against unfair competition. (2) Any act of competition contrary to honest practices in industrial or commercial matters constitutes an act of unfair competition.'
52 Hodgkinson & Corby Ltd v. Wards Mobility Services Ltd [1994] 1 WLR 1564 at 1569-1570 per Jacob J: 'I...begin by identifying what is not the law. There is no tort of copying. There is no tort of taking a man’s market or customers. Neither the market nor the customers are the plaintiff’s to own. There is no tort of making use of another’s goodwill as such. There is no tort of competition.'
54 Ibid. at 492-493: '...competition must remain free...A defendant...does no wrong by entering a market created by another and there competing with its creator. The line may be difficult to draw; but, unless it is drawn, competition will be stifled...'
trader is prevented from protecting their goodwill in a passing off action,\textsuperscript{55} thus restricting the level of competition in new entries of goodwill that can be established.\textsuperscript{56}

Such practice is outlawed in other jurisdictions such as Germany,\textsuperscript{57} and America,\textsuperscript{58} although, there have been calls for similar protection in Britain. Indeed, Mills comments that a fair system should be in place in order to provide fair competition between traders.\textsuperscript{59} The Unfair Commercial Practices Directive\textsuperscript{60} attempts to harmonise consumer protection laws in Europe as it requires Member States to abstain from unfair competition in the business-to-consumer field, to reduce the likelihood of consumer confusion among goods or services.\textsuperscript{61}

Howells\textsuperscript{62} comments that the implementation of the Directive has provided the UK with the means to simplify consumer law, although it remains to be seen if the level of protection afforded has been increased.\textsuperscript{63} Indeed, Andrew Gowers\textsuperscript{64} recommends that the measures provided in the Directive should be monitored

\textsuperscript{55} Gowers Review of Intellectual Property (London: The Stationary Office, 2006) p 100, para. 5.84.

\textsuperscript{56} Id; Cornish W.R. ‘Genevan Bootstraps’ Vol. 19 (7) (1997) European Intellectual Property Review 336-338 at 337: ‘...unfair competition liability...can all too easily become a weapon by which first entrants on to successful markets can engage in legalistic bullying of those who would subsequently seek to compete with them...’


\textsuperscript{58} International News Service v. Associated Press 248 US 215 (1918) at 258: “…The unfairness in competition which hitherto has been recognized by the law as a basis for relief, lay in the manner or means of conducting the business; and the manner...held legally unfair, involves either fraud or force or the doing of acts otherwise prohibited by law…”

\textsuperscript{59} Mills, B. 'Own label products and the "lookalike" phenomenon: a lack of trade dress and unfair competition protection?” Vol. 17 (3) (1995) European Intellectual Property Review 116-132 at 132: ‘...if businesses are to be made to compete in the ‘game’ of commerce then the rules of the ‘game’ must also ensure that the competitors compete fairly.’

\textsuperscript{60} Unfair Commercial Practices Directive 2005/29/EC.

\textsuperscript{61} Ibid. at Recital 8: ‘This Directive directly protects consumer economic interests from unfair business-to-consumer commercial practices. Thereby, it also indirectly protects legitimate businesses from their competitors who do not play by the rules in this Directive and thus guarantees fair competition in fields coordinated by it. It is understood that there are other commercial practices which, although not harming consumers, may hurt competitors and business customers. The Commission should carefully examine the need for Community action in the field of unfair competition beyond the remit of this Directive and, if necessary, make a legislative proposal to cover these other aspects of unfair competition.’


\textsuperscript{63} Ibid. at 193-194.

by the Government to prevent the notion of unfair competition and if found to be inadequate, then appropriate changes must be considered and implemented.65

It therefore appears that the current law adequately balances the needs of consumers and the interests of traders by allowing consumers to distinguish between goods or services provided by traders, due to the mark having to satisfy the requirement in terms of identifying origin. This, in turn, creates perceptions of quality among consumers, especially as the practice of comparative advertising, also allows consumers to make informed choices during purchase, whilst traders are able to protect their mark from infringement under s.10 Trade Marks Act 1994,66 and under the common law tort of Passing off. However, it appears that the promotion of fair competition is required in order to reduce the likelihood of consumer confusion with regard to 'look-a-like' packaging,67 as well as protecting the trader with goodwill that has been established after the look-a-like product has been made.

Innovators

With regard to Patent Law, s. 25 Patents Act 1977 gives the creator or inventor68 of an invention a monopoly right lasting 20 years provided that the information with regard to the invention has been disclosed,69 and clearly stated,70 with the information and invention based on a product or process.71

This monopoly provides the inventor with an incentive to innovate, by providing them with a right to prevent others from taking unfair advantage of their invention through selling (including offering for sale), making, importing or using the invention without the inventors consent.72 However, the fact that the

---

65 Ibid. at 100, para. 5.88.
66 Trade Marks Act 1994 s. 10.
68 Patents Act 1977 s. 7(3): "..."inventor" in relation to an invention means the actual deviser of the invention and "joint inventor" shall be construed accordingly."
69 Ibid. at s.25(1).
70 Ibid. at s. 14(5)(b).
71 Ibid. at s. 60(1)(b).
72 Ibid. at s. 60(1); Gowers Review of Intellectual Property (London: The Stationary Office, 2006) p 13, para. 1.11.
monopoly lasts for 20 years illustrates the existence of a high standard that must be satisfied in order to obtain protection. Indeed, to obtain a Patent the inventor must prove that the invention is 'new', possesses an 'inventive step' and is 'capable of industrial application', under s. 1(1) of the 1977 Act.\footnote{Patents Act 1977 s. 1(1).}

This monopoly right also accords with the Labour theory in the sense that it rewards the inventor for the time spent in developing the invention,\footnote{Locke, J., Two Treatises of Government (Cambridge: Cambridge University Press, 1988) p 287-288, para. 27; Chiron Corp v. Organon Teknika Ltd (No.10) [1995] FSR 325 at 332.} as well as inducing the inventor to disclose their discovery in order to be protected.\footnote{Id.} The monopoly right therefore encourages technical progress as the requirement of disclosure adds to the store of available knowledge.\footnote{Id.; Hickton's Patent Syndicate v. Patents and Machine Improvements Co Ltd (1909) 26 RPC 339.} Indeed, in Genentech Inc's Patent,\footnote{Genentech Inc's Patent [1987] RPC 553.} Whitford J observed that once people are informed of a process then a Patent may be employed as a reward of the discovery.\footnote{Ibid. at 566: '...It is trite law that you cannot patent a discovery, but if on the basis of that discovery you can tell people how it can be usefully employed, then a patentable invention may result. This...would be the case, even though once you have made the discovery, the way in which it can be usefully employed is obvious enough...'; Genentech Inc's Patent (Human Growth Hormone) [1989] RPC 147 at 240.}

This monopoly right, in addition to the Labour theory, also accords with the Utilitarian theory as favoured by Jeremy Bentham, who claimed that the interposition of the law in securing such benefits is considered to be a necessity in furthering economic interests.\footnote{Bentham, J., The Works of Jeremy Bentham, Vol. 3 (Edinburgh: John Bowring, 1843) p. 71: ‘...Without the assistance of the laws, the inventor would almost always be driven out of the market by his rival, who finding himself, without any expense, in possession of a discovery which has cost the inventor much time and expense, would he able to deprive him of all his desired advantages, by selling at a lower price. An exclusive privilege is of all rewards the best proportioned, the most natural, and the least burthensome. It produces an infinite effect, and it costs nothing...’} Indeed, such state of affairs prompted Aldous J in Chiron Corp v. Organon Teknika,\footnote{Chiron Corp v. Organon Teknika Ltd (No.10) [1995] FSR 325.} to comment that it is in the public interest for the inventors monopoly right to be enforced.\footnote{Ibid. at 333-334.}

\textit{Interpretation of Claims}

The English legal system has typically revolved around a literal approach to interpreting legal documents,\footnote{R v. City of London Court Judge and Payne [1892] 1 QB 273 at 290, per Lord Esher MR: ‘...If the words of an Act are clear, you must follow them, even though they lead to a manifest absurdity...’} and the same has been true with regard to construing Patent specifications and claims, when addressed to persons skilled in the art in which the patent is granted.\footnote{Electric & Musical Industries Ltd v. Lissen Ltd [1938] 4 All ER 221 at 224-225, per Lord Russell: ‘...The function of the claims is to define clearly and with precision the monopoly claimed, so that others may know the exact boundaries of the area within which they will be trespassers. Their primary object is to limit, and not to extend, the monopoly. What is not claimed is disclaimed...A patentee who describes an invention in the body of a specification obtains no monopoly unless it is claimed in the claims.’; Rodi & Wienenberger AG v. Henry Showell Ltd [1968] FSR 100 at 104.} Although, this ‘flagpost’ approach would sometimes lead to harsh results.\footnote{Zander, M., The Law-Making Process, 6th edn (Cambridge: Cambridge University Press, 2004) p145: ‘...the literal approach to interpretation ...is defeatist and lazy. The judge gives up the attempt to understand the document at the first attempt...It is the intellectual equivalent of deciding the case by tossing a coin...the literal approach is [therefore] always wrong...’} Indeed, in Van der Lely v. Bamfords Ltd,\footnote{Van der Lely NV v. Bamfords Ltd [1964] RPC 54.} it was identified that a 'pith and marrow' principle existed, whereby the variant of an infringing product was held not to infringe as the patent claim had expressly identified another integer as being essential to the invention claimed, even though the infringing integer performed the same function.\footnote{Id.}

Such a result therefore required a need for balancing the patentees protection, and the rights of third parties to determine the extent of the granted monopoly, to avoid any potentially infringing acts.\footnote{Turner, J. D. C. 'Purposive construction' (2001) Vol. 23 (2) European Intellectual Property Review 118: ‘...there [appeared] to have been...[the need for]...a major shift from strict to liberal construction.’}

The requirement of balance came to ahead in Catnic Components v. Hill & Smith,\footnote{Id.} where Lord Diplock adopted a Purposive approach to interpretation, which

\footnotesize
\begin{itemize}
  \item \footnote{R v. City of London Court Judge and Payne [1892] 1 QB 273 at 290, per Lord Esher MR: ‘...If the words of an Act are clear, you must follow them, even though they lead to a manifest absurdity...’}
  \item \footnote{Electric & Musical Industries Ltd v. Lissen Ltd [1938] 4 All ER 221 at 224-225, per Lord Russell: ‘...The function of the claims is to define clearly and with precision the monopoly claimed, so that others may know the exact boundaries of the area within which they will be trespassers. Their primary object is to limit, and not to extend, the monopoly. What is not claimed is disclaimed...A patentee who describes an invention in the body of a specification obtains no monopoly unless it is claimed in the claims.’; Rodi & Wienenberger AG v. Henry Showell Ltd [1968] FSR 100 at 104.}
  \item \footnote{Zander, M., The Law-Making Process, 6th edn (Cambridge: Cambridge University Press, 2004) p145: ‘...the literal approach to interpretation ...is defeatist and lazy. The judge gives up the attempt to understand the document at the first attempt...It is the intellectual equivalent of deciding the case by tossing a coin...the literal approach is [therefore] always wrong...’}
  \item \footnote{Van der Lely NV v. Bamfords Ltd [1964] RPC 54.}
  \item \footnote{Id.}
  \item \footnote{Turner, J. D. C. 'Purposive construction' (2001) Vol. 23 (2) European Intellectual Property Review 118: ‘...there [appeared] to have been...[the need for]...a major shift from strict to liberal construction.’}
\end{itemize}
involved observing the function of the Patented invention in order discover
whether the same result was being achieved by the alleged infringer.90
Consequently, it was decided that the defendants lintel did infringe as its
verticality was essential to its function.91

This Purposive approach to interpretation was subsequently adopted under Art.
69 of the European Patent Convention (EPC) in order to provide fair protection
to the patentee,92 and is adopted in English law through s. 125(3) Patents Act
1977,93 for the purposes of s. 125(1) of the Act.94

The purposive approach to interpretation however, is not universally accepted.
Indeed, following the decision in Kastner v. Rizla,95 where it was found that the
defendants similar process for cutting cigarette papers infringed the claimants
process,96 Oliver,97 observed that the method of interpretation had come close to
protecting virtually any functional equivalent of the patented invention as
illustrated in the case by Aldous LJ, who had taken a wide view as to the source

90 Ibid. at 65-66: ‘...A patent specification should be given a purposive construction rather than a
purely literal one derived from applying to it the kind of meticulous verbal analysis in which lawyers
are too often tempted by their training to indulge...’; Magor and St Mellons Rural District
Council v. Newport Corporation [1950] 2 All ER 1226 at 1236, per Denning LJ: ‘...We sit here to
find out the intention of Parliament...and carry it out, and we do this better by filling in the gaps
and making sense of the enactment than by opening it up to destructive analysis.’; Heydons Case
(1586) 74 ER 67.
should not be interpreted as meaning that the extent of the protection conferred by a European
patent is to be understood as that defined by the strict, literal meaning of the wording used in the
claims, the description and drawings being employed only for the purpose of resolving an ambiguity
found in the claims. Nor should it be taken to mean that the claims serve only as a guideline and
that the actual protection conferred may extend to what, from a consideration of the description
and drawings by a person skilled in the art, the patent proprietor has contemplated. On the
contrary, it is to be interpreted as defining a position between these extremes which combines a fair
protection for the patent proprietor with a reasonable degree of legal certainty for third parties.’
Patent Convention...shall, as for the time being in force, apply for the purposes of subsection (1)
above as it applies for the purposes of that Article.’
94 Ibid. at s. 125(1); Pharmacia Corp v. Merck & Co Inc [2002] RPC 41.
96 Ibid.
Intellectual Property Review 28-34.
of the invention, despite the claimant having a specific process of achieving the end result.\(^{98}\)

Turner\(^{99}\) has commented that the limited guidance offered in the EPC fails to form the basis of future development, as the British courts have failed to contribute to developing the guidance, by simply following the approach laid down in *Catnic*.\(^{100}\) He therefore suggests that the Catnic principle should be discarded, despite the approach being favoured by Balcombe LJ in *Kastner* who was not prepared to abandon the many years of case law following the *Catnic* approach, in favour of a new approach.\(^{101}\) Indeed, Lord Walker in *Kirin-Amgen v. Hoechst Marion Roussel*,\(^{102}\) commented that the principle in *Catnic* may be difficult to apply in areas of fast-developing technology which is becoming increasingly common in the Patent field,\(^{103}\) and therefore implies that a new approach is needed.

In *PLG Research v. Ardon*,\(^{104}\) Neill LJ doubted the extent of protection under the 1977 Act as he maintained, *obiter*, that an element of flexibility existed allowing the English courts to enact a literal approach.\(^{105}\) Indeed, Watson and Karet\(^{106}\) acknowledged that the literal approach offers a higher degree of certainty, as a potential competitor would be able to determine the extent to which their product or process would not infringe the established invention,\(^{107}\) as well as reducing the level of uncertainty due to the requirement of precise and

\(^{98}\) Id.; *Kastner v. Rizla Ltd (No.1)* [1995] RPC 585; In general criticism, Lord Simonds observed that the use of the purposive approach appeared ‘...to be a naked usurpation of the legislative function under the thin disguise of interpretation...If a gap is disclosed, the remedy lies in an amending Act.’ *Magor and St Mellons Rural District Council v. Newport Corporation* [1952] AC 189 at 191.


\(^{101}\) Turner, J. D. C. ‘Purposive construction: seven reasons why Catnic is wrong’ Vol. 21 (11) (1999) *European Intellectual Property Review* 531-536 at 535-536; *Kastner v. Rizla* [1995] RPC 585 at 605; ‘...there can be no basis for discardimg...14 years of case law...If the Catnic approach is inconsistent with the Protocol...then it is for the House of Lords...to say so.’

\(^{102}\) *Kirin-Amgen Inc v. Hoechst Marion Roussel Ltd* [2004] UKHL 46.

\(^{103}\) *Ibid.* at 139.


\(^{105}\) *Ibid.* at 133.


appropriate language used in the Patent claims. Consequently, Neill LJ opined that Lord Diplock’s approach to purposive construction should be cast into legal history.

Lord Hoffmann in Kirin-Amgen, disagreed, and compared this argument to the Caliph Omar’s justification for burning the library of Alexandria. Indeed, Cole was of opinion that precision when drafting patent claims cannot always be achieved, especially when trying to describe undeveloped and complicated technological areas. Lord Hoffmann therefore acknowledged that the principle in Catnic was consistent with Protocol on Art. 69 as it was aimed at giving the patentee a full monopoly.

Despite this full monopoly, there have been differences of interpretation of Art. 69 of the Protocol in different jurisdictions. In Improver v. Remington Consumer Products, the Court of Appeal in Hong Kong, which was subject to British rule at the time and therefore subject to the Patents Act 1977, found that the claimants hair removal device did not infringe the defendants, as a person skilled in the art in which the patent was granted, may have understood that the patentee expected strict compliance as a necessary requirement of the invention. Even though, the variant did not have an effect on the function of the invention, as well as being obvious to a person skilled in the art from the date the patent was published.

---

108 Id.; Conoco Speciality Products v. Merpro Montassa Ltd (No. 3) 1992 SLT 444.
111 Ibid. at para. 46.
113 Kirin-Amgen Inc v. Hoechst Marion Roussel Ltd [2004] UKHL 46, para. 48: ‘The Catnic principle of construction is therefore in my opinion precisely in accordance with the Protocol. It is intended to give the patentee the full extent, but not more than the full extent, of the monopoly which a reasonable person skilled in the art, reading the claims in context, would think he was intending to claim...’
117 Ibid. at 189.
118 Id.
119 Id.
Although, when the same case was conducted in Germany through a parallel action, the German court held that infringement had occurred.\textsuperscript{120} Jacob,\textsuperscript{121} confessed that he was unsure as to the justifications for reaching this decision, but suspected that the reasons related to the cultural differences in the common law.\textsuperscript{122}

However, in \textit{Kirin-Amgen}, it was observed that the German courts were moving closer to the approach of interpretation taken by the Hong Kong Court of Appeal to the extent that the German system was similar to the approach in \textit{Catnic}.\textsuperscript{123} Indeed, Aldous LJ, observed that the \textit{Improver} approach accorded with the Protocol on Art. 69.\textsuperscript{124}

The Protocol, the principle laid down in \textit{Catnic}, and the \textit{Improver} tests therefore attempt to strike a balance between the protection of the patentees invention and certainty for the public interest in order to determine the extent of the granted monopoly, when viewing the patent specification, to avoid infringement when engaging in a related activity.\textsuperscript{125} Indeed, in \textit{Merck v. Generics (UK)}.\textsuperscript{126} Laddie J commented that:

\begin{quote}
"The purpose of a patent is to convey to the public what the patentee considers to be his invention and what monopoly he has chosen to obtain...the patentee must be taken to know the framework of form and purpose when he drafts his patent. It is his duty to communicate his invention and his assertion of monopoly to the public in language it will understand..."\textsuperscript{127}
\end{quote}

\textsuperscript{120} Re Formstein [1991] RPC 597.
\textsuperscript{122} Ibid. at 313.
\textsuperscript{126} Merck & Co Inc v. Generics (UK) Ltd [2003] EWHC 2842 (Pat).
\textsuperscript{127} Ibid. at para. 38.
In AssiDoman Multipack v. Mead Corp,\textsuperscript{128} Aldous J commented that he could not think of a better way in aiming to achieve consistent decisions in domestic and community cases.\textsuperscript{129}

**Public Interest**

The term 'Copyright' directly describes its function, as it aimed at protecting the author from unauthorised reproductions or copies of their work.\textsuperscript{130} Copyright therefore protects the 'judgement, skill or labour' that has been put into the expression of an idea\textsuperscript{131} by the person who created the work.\textsuperscript{132} Copyright therefore does not protect the intangible idea itself.\textsuperscript{133}

This protection is awarded once the work has been published in the public domain with the consent of the copyright owner, otherwise an action for infringement may arise.\textsuperscript{134} Indeed, without legal protection many investments

---

\textsuperscript{128} AssiDoman Multipack v. Mead Corp [1995] FSR 225.

\textsuperscript{129} Ibid. at 236-237: ‘...For myself, I would be loath to discard 14 years of case law unless it is certain that “purposive” construction is not the correct approach under the Act...I conclude that the correct approach to construction under the Patents Act 1977 is “purposive” construction...I have been unable to think of any better guidance which hopefully will result in consistent decisions between the courts of this country and those of other parties to the Convention...’; Lord Hoffmann described this finding as ‘masterly’: Kirin-Amgen Inc v Hoechst Marion Roussel Ltd [2004] UKHL 46, para. 46; The test has also been adopted in other jurisdictions including Canada: Eli Lilly & Co v. Novopharm Ltd [1996] RPC 1.

\textsuperscript{130} Ladbroke (Football) Ltd v William Hill (Football) Ltd [1964] 1 WLR 273 at 291 per Lord Pearce: ‘...the work should not be copied but should originate from the author.’

\textsuperscript{131} Cala Homes (South) Ltd v. Alfred McAlpine Homes East Ltd (No.2) (1995) IP & T Digest 18 at per Laddie J: ‘...to have regard merely to who pushed the pen is too narrow a view of authorship. What is protected by copyright in a drawing or a literary work is more than just the skill of making marks on paper or some other medium. It is both the words or lines and the skill and effort involved in creating, selecting or gathering together the detailed concepts, data or emotions which those words or lines have fixed in some tangible form which is protected. It is wrong to think that only the person who carries out the mechanical act of fixation is an author...’; Sawkins v. Hyperion Records Ltd [2005] EWCA Civ 565; British Leyland Motor Corp Ltd v. Armstrong Patents Co Ltd [1986] AC 577.

\textsuperscript{132} Copyright, Designs and Patents Act (1988) s. 9(1); Berne Convention for the Protection of Literary and Artistic Works 1886 Art. 6bis: ‘(1)...the author shall have the right to claim authorship of the work and to object to any distortion, mutilation or other modification of, or other derogatory action in relation to, the said work, which would be prejudicial to his honor or reputation.’

\textsuperscript{133} University of London Press Ltd v. University Tutorial Press Ltd [1916] 2 Ch 601 at 608, per Peterson J: ‘...Copyright Acts are not concerned with the originality of ideas, but with the expression of thought...’

\textsuperscript{134} Copyright, Designs and Patents Act 1988 ss. 153-156.
would decrease as many people would be able to copy the work and sell it at a cheaper rate than the original author of the work. It is therefore in the public interest that legal protection is given to the authors work, as this will produce a stronger economy through further investment. Indeed, in Ellis v. Home Office, Morris LJ observed that it is in the public interest that justice is seen to be done.

This illustrates the operation of the Utilitarianism theory, associated with Jeremy Bentham in the sense that the public interest as a whole overrides the needs of the individual. Indeed, Bergh has observed that the general principle of granting legislation is to defend the public good. Consequently, the right between protecting the authors’ interests in publication of their work, and the public interest in accessing the work is required to be balanced, as copyright protects tangible ideas which are considered to be of some economic value.

Fair Dealing

The 1988 Act allows a variety of acts that can be undertaken by the public which will not infringe the authors work, provided that the use of the work is ‘fair’ for the purpose it is used, thereby balancing the public interest in gaining access to, and using the copyright work, subject to the copyright owners

136 Id.
138 Ibid. at 147.
139 Bentham, J., The Principles of Morals and Legislation (London: Prometheus Books, 1988) p 2, para. 3: '...utility is meant...to prevent the happening of mischief, pain, evil, or unhappiness to the party whose interest is considered: if that party be the community in general, then the happiness of the community...'
141 Ibid. at 20: '...Since the value of the work to society is not well reflected in the value of the work to the private owner, legal devices are needed to solve the public good problem...'
142 University of London Press Ltd v. University Tutorial Press Ltd [1916] 2 Ch 601 at 610 per Peterson J: '...there remains the rough practical test that what is worth copying is prima facie worth protecting...'
143 These are referred to as 'Permitted Acts': Copyright, Designs and Patents Act 1988 ss. 29-76.
144 Beloff v. Pressdram Ltd [1973] 1 All ER 241 at 262, per Ungoed-Thomas J: '...It is fair dealing directed to and consequently limited to and to be judged in relation to the approved purposes. It is dealing which is fair for the approved purposes and not dealing which might be fair for some other purpose or fair in general...'; Hubbard v. Vosper [1972] 2 QB 84 at 94.
rights.\textsuperscript{145} The 'permitted acts' therefore act as defences to an infringement action.\textsuperscript{146}

However, unlike the American system, the defences of 'fair use' apply only in certain situations, namely 'Research and private study'\textsuperscript{147} and in circumstances involving 'Criticism, review and news reporting',\textsuperscript{148} where sufficient acknowledgement of the title and author has been made.\textsuperscript{149} These defences are therefore not applicable in a general sense,\textsuperscript{150} despite being required to be applied liberally.\textsuperscript{151} Indeed, writing extra-judicially, Laddie J described that this constraint has created a 'rigid' rule of application.\textsuperscript{152}

When a question of infringement arises, the courts tend to focus on the quality rather than the quantity of the alleged infringement,\textsuperscript{153} by observing whether it falls within 'fair use' exceptions, as well as looking at the motives and purposes behind the publication,\textsuperscript{154} in order to determine whether the act is 'fair',\textsuperscript{155} on an objective basis.\textsuperscript{156} Although, this creates uncertainty as cases of infringement tend to be decided on an incremental basis. Indeed, in \textit{Hubbard v. Vosper},\textsuperscript{157} it

\textsuperscript{145} CCH Canadian Ltd v. Law Society of Upper Canada [2004] 5 LRC 428 at 446.

\textsuperscript{146} \textit{Copyright, Designs and Patents Act 1988 s. 28(1)}: 'The provisions of...[Chapter 3] specify acts which may be done in relation to copyright works notwithstanding the subsistence of copyright; they relate only to the question of infringement of copyright and do not affect any other right or obligation restricting the doing of any of the specified acts.'

\textsuperscript{147} \textit{Ibid. at s. 29}; \textit{Silitoe v. McGraw-Hill Book Co (UK) Ltd} [1983] FSR 545.

\textsuperscript{148} \textit{Copyright, Designs and Patents Act 1988 s. 30};

\textsuperscript{149} \textit{Ibid. at s. 178}: '..sufficient acknowledgement' means an acknowledgement identifying the work in question by its title or other description, and identifying the author unless— (a) in the case of a published work, it is published anonymously; (b) in the case of an unpublished work, it is not possible for a person to ascertain the identity of the author by reasonable inquiry...'


\textsuperscript{151} \textit{Pro Sieben Media AG v. Carlton UK Television Ltd} [1999] 1 WLR 605 at 614 \textit{per} Hoffmann LJ: ‘“Criticism or review” and "reporting current events" are expressions of wide and indefinite scope. Any attempt to plot their precise boundaries is doomed to failure. They are expressions which should be interpreted liberally...’


\textsuperscript{153} Hawkes & Son (London) Ltd v. Paramount Film Service Ltd [1934] Ch 593 at 606.

\textsuperscript{154} \textit{Pro Sieben Media AG v. Carlton UK Television Ltd} [1999] 1 WLR 605.

\textsuperscript{155} \textit{Hubbard v. Vosper} [1972] 2 QB 84 at 94.

\textsuperscript{156} \textit{Pro Sieben Media AG v. Carlton UK Television Ltd} [1999] 1 WLR 605 at 614 \textit{per} Hoffmann LJ: ‘... It is not necessary for the court to put itself in the shoes of the infringer of the copyright in order to decide whether the offending piece was published for the purposes of criticism or review...’; \textit{Hyde Park Residence Ltd v. Yelland} [2001] Ch 143 at 154-155, para. 21.

\textsuperscript{157} \textit{Hubbard v. Vosper} [1972] 2 QB 84.
was observed that fairness must be determined on a matter of impression, as 'fairness' is impossible to define.\textsuperscript{158}

Colston\textsuperscript{159} has commented that this rigidity has created an imbalance based on judicial failure to take into account the general fairness in disseminating copyright works.\textsuperscript{160} Indeed, in Sillitoe v. McGraw-Hill Book,\textsuperscript{161} it was acknowledged that fair use with regard to research and private study largely covered situations where students copied work for their own use, but not situations where the copying was to be distributed among other students.\textsuperscript{162}

Such decisions have prompted Burrell\textsuperscript{163} to argue that the restrictive arrangement of the 'fair use' exception, has allowed the judges to dismiss the public interest, in circumstances falling outside of the exceptions.\textsuperscript{164} As a result, Burrell suggests that the judges should adopt a similar approach to the American system, in order to take account of the public interest in protecting the copyright owners work from illegal copies of their work being made.\textsuperscript{165}

Indeed, Andrew Gowers has suggested that the defence of private copying for research purposes should be applied to all forms work, including films and sound recordings, as this exclusion is felt to hinder students academic studies.\textsuperscript{166}

Indeed, the Intellectual Property Office (IPO) has highlighted that extending the defence would allow a greater range of material to be accessed by the public,

\textsuperscript{158} Ibid. at 94.
\textsuperscript{160} Ibid. at 106.
\textsuperscript{162} Ibid. at 558, per Davies J: 'The onus of showing that an exception applies is on the defendants. Mr. Jeffs contendd that [the defence] is widely drawn and not limited to the actual student, so that if a dealing is fair and for the purposes of private study...[it]...applies whether the private study in mind is one's own or that of somebody else...he said, the dealing was for the purpose of private study by the examinees who would acquire the notes. I do not accept that argument...'
\textsuperscript{163} Burrell, R. 'Reining in copyright law: is fair use the answer?' (4) (2001) Intellectual Property Quarterly 361-388.
\textsuperscript{164} Ibid. at 388.
\textsuperscript{165} Id.
which would increase public knowledge, and improve students and researchers skills with regard to using new technology.

Although, in considering Gowers recommendations for reform, in 2007 the IPO felt that it was unnecessary to implement the suggestions, as the cases that have been decided, provide sufficient assistance as to what is considered to be 'fair dealing'.

In the second stage of the consultation in 2009 however, the IPO felt that it was time to extend the defence based on misunderstanding from the responses received by universities and libraries with regard to the circumstances in which the defence permitted lawful copying. Consequently, the IPO feels that the defence should be extended, but on condition that its users are students of an educational establishment, in order to avoid the risk of unauthorised copying for 'entertainment purposes'. The IPO propose that the changes are due to come into force sometime in 2010 under a Statutory Instrument, although, only time will tell if the propositions have had any effect upon the public interest.

Griffiths acknowledges that in spite of the restrictive approach taken by the courts, overall, an element of consistency exists in favour of the protection of freedom of movement, as it the incremental approach taken by the courts have sought ways to avoid the limitations of the defences. Consequently, he comments that the obiter statement made in Sillitoe, based on a recognition of an

---

168 Ibid. at para. 136.
169 Ibid. at 26, para. 160.
170 Ibid. at 25, para. 158.
172 Ibid. at 28, para. 186.
173 Ibid. at p 35.
174 Ibid. at p 47-53.
176 Ibid. at 175.
177 Id.
authors claims to be regarded as being 'sufficient acknowledgment', must now be treated as wrong.\textsuperscript{178}

It therefore appears that in trying to balance the needs of both the copyright owner and the public interest, further measures are required to satisfy the public interest in order to avoid infringement, when gaining access to the copyright owners work, despite the consistency among the courts in placing limits on the fair dealing defences.

\textit{Conclusion}

In conclusion, it appears that, despite the criticisms that have been made, there exists an adequate balance of intellectual property rights which is protected by the law. Indeed, Andrew Gowers acknowledged that the current system overall, is satisfactory.\textsuperscript{179} However, the criticisms that have been made, serve as a suggestion for improving the current system, in order to provide greater protection to inventors to increase incentives to innovate,\textsuperscript{180} which will enhance consumer choice during purchases due to the variety of goods or services offered, and will further the public interest in providing greater investments to strengthen the economy.\textsuperscript{181} Indeed, Gowers has commented that a stronger balance is needed to compete in the 'global knowledge-based economy'.\textsuperscript{182}

\hspace{1cm} ANDREW PELLINGTON

\textsuperscript{180} \textit{Ibid.} at 119, para. 7.1.
\textsuperscript{181} \textit{Ibid.} at 119.
\textsuperscript{182} \textit{Ibid.} at para. 7.1.
Bibliography

Books

Cases
Aristoc Ltd v. Rysta Ltd [1945] AC 68.
Arsenal Football Club plc v. Reed [2001] 2 CMLR 481.
(Case 206/01) Arsenal Football Club Plc v. Reed [2003] Ch 454.
Commissioners of Inland Revenue Appellants v. Muller [1901] AC 217.
Conoco Speciality Products v. Merpro Montassa Ltd (No. 3) 1992 SLT 444.
Draper v. Trist [1939] 3 All ER 513.
Hawkes & Son (London) Ltd v. Paramount Film Service Ltd [1934] Ch 593.
Heydons Case (1586) 74 ER 67.
Hodgkinson & Corby Ltd v. Wards Mobility Services Ltd [1994] 1 WLR 1564.
Ladbroke (Football) Ltd. v William Hill (Football) Ltd [1964] 1 WLR 273.
Perry v. Truefitt (1842) 49 ER 749.
Re Coca-Cola’s Trademark Application; Coca-Cola Ltd v. Pike (No. 2472042) [2010] WL 666343.
Victoria Park Racing Co v. Taylor (1937) 58 CLR 479.
Journals


**Legislation**
Berne Convention for the Protection of Literary and Artistic Works 1886.
Comparative Advertising Directive 2006/114/EC.
Control of Misleading Advertisements (Amendment) Regulations 2000.
Consumer Protection Act 1987 c. 43.
Copyright Act 1976 (US).
Council Regulation (EC) 40 /94.
Forgery and Counterfeiting Act 1981, c. 45.
German Act Against Unfair Competition 1909: Gesetz gegen den unlauteren wettbewerb-UWG.
Paris Convention for the Protection of Industrial Property 1883.
Patents Act 1977 c. 37.
Universal Declaration of Human Rights 1948.

**Reports**
Bulletin of the European Communities, Supplement 8/76.

**Websites**
Overriding Interests under the Land Registration Act 2002

MATTHEW ASTLEY

The cases of Thomas v Clydesdale Bank,1 Thompson v Foy2 and Link Lending v Bustard3 are the first reported cases which directly consider the operation of paragraph 2 of Schedule 3 to the Land Registration Act 2002 (‘the LRA 2002’). In particular, these cases focus on the interpretation of the phrase ‘actual occupation’ under that paragraph. Of particular interest is how this term will be interpreted in the post-LRA 2002 era. Indeed the cases have already provoked academic comment on this point.4 However, whilst it is necessary to consider actual occupation in some detail, this article is principally concerned with the exception under paragraph 2(c)(i). This exception was considered in both Thomas and (albeit less significantly) in Thompson. In Link Lending the Appellant chose not to rely upon the exception and so it was mentioned by the Court of Appeal only in passing.5

In light of Thomas (and also of Thompson), this article is concerned with the following issues: the precise wording and interpretation of paragraph 2(c)(i); how and to what extent the obviousness test under paragraph 2(c)(i) differs from the actual occupation test; whether paragraph 2(c)(i) places a duty on disponees to conduct an actual inspection of the property, and if so the extent and scope of that duty and what it entails; and whether the operation of the obviousness test is, or is apt to be, particularly problematic in practice. The case of Thomas will be used as the primary basis for the exploration of these issues, and regard will be had to Thompson and Link Lending where appropriate.

---

2 [2009] EWHC 1076 (Ch).
5 Para 10.
The statutory framework
Schedule 3 of the LRA 2002 enumerates certain unregistered interests which override registered dispositions of land.\(^6\) Paragraph 2 of that Schedule provides that, subject to certain exceptions, an interest belonging at the time of the disposition to a person in actual occupation, so far as relating to land of which he is in actual occupation, will override a registered disposition.

One of the exceptions to this rule is provided for in paragraph 2(c). Under paragraph 2(c)(i) the interest of a person in actual occupation will not override a registered disposition if that person’s occupation ‘would not have been obvious on a reasonably careful inspection of the land at the time of the disposition’.

Further, under paragraph 2(c)(ii), the interest of a person in actual occupation will not override if it is an interest ‘of which the person to whom the disposition is made does not have actual knowledge’ at the time of the disposition. In practice this operates as an exception to the requirement under paragraph 2(c)(i).\(^7\)

Prior to the LRA 2002, the overriding interests of those in actual occupation was governed by section 70(1)(g) LRA 1925. The section protected the ‘rights of every person in actual occupation of the land... save where enquiry is made of such person and the rights are not disclosed.’ Paragraph 2(b) of Schedule 3 effectively reformulates the latter part of section 70(1)(g) as regards the disclosure of rights.\(^8\) It places a burden on disponees to enquire about the interests of each occupier and a reciprocal burden on occupiers to disclose any interests belonging to them where it is reasonable to expect them to do so.

Section 70(1)(g) was ‘notorious and much-litigated’.\(^9\) The litigation focussed, among other things, on the definition of and what constitutes actual occupation (and as Thomas, Thompson and Link Lending illustrate, this remains a

\(^6\) See too ss 29(1) and 29(2)(a)(ii).


contentious issue). Particular problems included whether interests were binding on disponees despite being difficult to discover.\(^\text{10}\) Accordingly the LRA 2002 introduced a new element to what had been section 70(1)(g) in the form of paragraph 2(c) of Schedule 3. The rationale of this new provision, which provides one example of the way in which the LRA 2002 aims to limit the scope of overriding interests as far as possible, is ‘to protect buyers and other registered disponees for valuable consideration where the fact of occupation is neither subjectively known to them nor readily ascertainable.’\(^\text{11}\)

Nevertheless, particularly problematic is the situation where a person informally acquires an interest in land and the subsequent disposition of that land – through the creation of a charge, for example – is to a person who is unaware that he is dealing with a sole trustee. At the heart of these issues there exists a tension between vulnerable occupiers and unsuspecting disponees, and between the need to guarantee residential security, on the one hand, and the need for an efficient system of lending and conveyancing, on the other.\(^\text{12}\)

**Thomas v Clydesdale Bank**

In January 2010 the Claimant and Respondent in the appeal, Clydesdale Bank Plc (‘the Bank’), was granted possession of a property in Leeds. The property was purchased in March 2006 by the First Defendant, Mr Burtenshaw, with a mortgage of £550,000 from the Bank of Scotland (the purchase price was £650,000). Before that juncture Mr Burtenshaw had formed a relationship with the Second Defendant and Appellant, Ms Thomas, with whom he lived in the latter’s house.

In about May 2006 Mr Burtenshaw approached the Bank. He sought a loan which would allow him to renovate the new property. The Bank offered him a secured bridging loan facility for £750,000 (allowing £200,000 for renovation works) and on 27 July 2006 a Deed of Mortgage of the property was executed by Mr

---

\(^{10}\) Sparkes, ‘The Discoverability of Occupiers of Registered Land’ (1989) *Conv* 342; Martin, ‘Section 70(1)(g) and the Vendor’s Spouse’ (1980) *Conv* 361.

\(^{11}\) Law Comm No 271, para 8.62.

\(^{12}\) Gray and Gray, *supra*, para 8.2.97.
Burtenshaw to secure all monies owing to the Bank, including further advances. Mr Burtenshaw, Ms Thomas and their children moved into the new property on 30 September that year.

Mr Burtenshaw fell into financial difficulties towards the end of 2007, and in May 2008 the Bank demanded repayment of the money owing on the mortgage account. Proceedings were commenced against Mr Burtenshaw for the outstanding monies in June 2008. Ms Thomas contended that she had a beneficial equitable interest in the property and she was joined into the proceedings later that year. Ms Thomas contended, specifically, that her beneficial interest arose by way of a common intention constructive trust prior to the date when the Deed of Mortgage was executed, and that her interest overrode that registered disposition.

The trial date was set for January 2010. However, Ms Thomas was unable to attend due to a depressive illness. (Mr Burtenshaw had no personal interest in the outcome by this juncture as he had been adjudged bankrupt.) She therefore applied for an adjournment, but the judge decided to proceed in her absence (in so doing he had CPR 39.3(3) in mind) and found in favour of the Bank, granting possession. Ms Thomas sought to have the judgment set aside under CPR 39.3(3), but the application was refused, and Ms Thomas then appealed against that refusal.

In her application under CPR 39.3(3) to have the original decision set aside, Ms Thomas established that she had acted promptly in making the application and, furthermore, that she had good reason for not attending the trial: CPR 39.3(5)(a)-(b). However, she failed to establish that she had a reasonable prospect of success at trial as required by CPR 39.3(5)(c) and accordingly appealed on this point.

The appeal itself centred on all the provisions of paragraph 2 of Schedule 3 (and in particular paragraph 2(c)). Ramsey J concluded that Ms Thomas had reasonable prospects of establishing (a) that she was in actual occupation; and
(b) that on a reasonably careful inspection her occupation would have been obvious. He further concluded that she had reasonable prospects of succeeding in showing that the Bank had actual knowledge of the facts giving rise to the alleged interest. The appeal was therefore allowed.

**Actual occupation under paragraph 2(c)**

In *Thomas* the property was not being used as a residence on the date of the disposition, namely 27 July 2006. Indeed it was not used for that purpose until Ms Thomas moved into the property, together with Mr Burtenshaw and her children, on 30 September. However, there was evidence to show that builders and interior designers were there on 27 July. This appears not to have been disputed by the Bank.¹³

Ms Thomas contended that the builders and interior designers were working in the property both on her behalf and on behalf of Mr Burtenshaw. This, it was submitted, combined with her attendance at the property at least every other day and her intention to reside there, established that she had a reasonable prospect of success on the issue of actual occupation.¹⁴ The Bank, on the other hand, contended that the presence of the builders and interior designers did not constitute occupation by Ms Thomas because she did not employ them. It was submitted that when she visited the property, she did so only as an agent of Mr Burtenshaw, and that her intention to return was immaterial.¹⁵ The requisite permanence and continuity, the Bank submitted, was absent.¹⁶

In deciding in favour of Ms Thomas, Ramsey J had regard to the decision in *Lloyds Bank v Rosset*¹⁷ and concluded that Ms Thomas's occupation was 'of the nature, and to the extent, that one would expect of an occupier having regard to the then state of the property in the state of renovation for residential use.'¹⁸ He

---

¹³ Para 19.
¹⁴ Para 16.
¹⁵ Para 19.
¹⁶ *Ibid*.
¹⁸ Para 33.
also applied the facts of the case to the factors listed as relevant by Mummery LJ in *Link Lending*.\(^1\) He stated that Ms Thomas

has reasonable prospects of establishing that there was a degree of permanence and continuity in her presence; that her intention and wish was that she should reside permanently at the property and that her presence was sufficient for the nature of the property in the course of renovation.\(^2\)

Also relevant, according to Mummery LJ, is the length of any absence and the reason for it.\(^3\) In *Link Lending* the Respondent had been in a residential care home, owing to psychiatric problems, for over two years. She was not, therefore, personally present in the property concerned during this period. The Appellant wanted to take possession. However, in dismissing the appeal, Mummery LJ noted (among other things) that 'it was her furnished home and the only place to which she genuinely wanted to return', that 'she continued to visit the Property because she still considered it her home', and that 'she was in the process of making an application to the Mental Health Review Tribunal in order to be allowed to return home.'\(^4\) Accordingly, the case was not one of 'mere fleeting presence' and the conclusion at first instance

was supported by evidence of a sufficient degree of continuity and permanence of occupation, of involuntary residence elsewhere, which was satisfactorily explained by objective reasons, and of a persistent intention to return home when possible, as manifested by her regular visits to the Property.\(^5\)

Under the LRA 1925, the courts adopted a literal interpretation of actual occupation; indeed Lord Wilberforce stated in *Williams & Glyn’s Bank v Boland*

\(^{19}\) *Link Lending*, para 27.
\(^{20}\) *Thomas*, para 33.
\(^{21}\) *Link Lending*, para 27.
\(^{22}\) Para 26.
\(^{23}\) Para 30.
that the words ‘actual occupation’ were ‘ordinary words of plain English, and should... be interpreted as such.’\(^\text{24}\) Thus, as Bogusz argues, what is noteworthy about the judgments in both *Thomas* and *Link Lending* is the apparent willingness on the part of the judiciary to ‘move beyond a purely objective assessment of the factual circumstances’ and ‘to incorporate the subjective intentions of a party when determining the existence of actual occupation.’\(^\text{25}\) The judgments arguably represent a broader, reflexive and more contextual approach to actual occupation than hitherto seen.\(^\text{26}\)

**The obviousness test under paragraph 2(c)(i)**

In *Thomas* Ramsey J gave much briefer consideration to the exception under paragraph 2(c)(i) than he did to the test for actual occupation. Nevertheless, this aspect of the judgment is helpful and important, for it sheds light on how the courts might interpret paragraph 2(c)(i) and in particular on how the question posed by the provision differs from and relates to the determination of whether a person is in actual occupation.

Ramsey J pinpointed the question as being whether Ms Thomas had reasonable prospects of establishing that the degree of occupation which he found would have been obvious on a reasonably careful inspection of the property.\(^\text{27}\) He concluded that Ms Thomas would in fact have reasonable prospects of establishing that the person inspecting would have been aware that she was visiting the property, and that builders and interior designers were working there.\(^\text{28}\)

As noted, under paragraph 2(c)(i) the relevant consideration is whether the occupation of the relevant person ‘would not have been obvious on a reasonably careful inspection of the land’. In *Thomas* Ramsey J described the phrase

\(^{24}\) [1980] 2 All ER 408, 412-413.
\(^{25}\) Bogusz, *supra*, 279.
\(^{26}\) *Ibid*, 284.
\(^{27}\) Para 37.
\(^{28}\) Paras 38 and 40.
‘reasonably careful inspection’ as ‘objective’ and stated that he found it difficult to read into it

a requirement that the person inspecting would have any particular knowledge or that, in the absence of any express provision, the term “inspection” would also require the person inspecting to make reasonable enquiries. On that basis, it is the visible signs of occupation which have to be obvious on inspection.29

Ramsey J was therefore unequivocal in relation to the focus being on visible signs of occupation, a point he also made a couple of paragraphs earlier on in his judgment.30 In doing so he highlighted the difference between the focus of paragraph 2(c)(i) and the focus of, and what is entailed by, the test for actual occupation. For, with regard to the latter test, during his consideration of paragraph 2(c)(i), Ramsey J said

It is clear… that, in order to determine whether somebody is in actual occupation it is necessary to determine not only matters which would be obvious on inspection but matters which would require enquiry to ascertain them. That includes such things as the permanence and continuity of the presence of the person concerned, the intentions and wishes of that person and the personal circumstances of the person concerned.

An inspection which only required something to be visible might show that somebody was staying at the property but without enquiry it would not be known whether they were there for an hour, a day, a month, or had some more permanent continuous presence to give rise to the required occupation.31

29 Para 40.
30 Para 38.
31 Paras 38-39.
It is therefore clear that in determining whether a person’s occupation would have been obvious, there is no requirement for a disponee to make reasonable enquiries. By contrast, in order to determine whether somebody is in actual occupation, it may be necessary to consider those matters which would require enquiry to ascertain them, such as those matters which have recently been considered by the courts as part of a reflexive and more contextual approach in relation to this issue.

Comment

The approach of Ramsey J in *Thomas* in relation to paragraph 2(c)(i), namely that there is no requirement for a disponee to make reasonable enquiries, is consistent with the interpretation of the exception by Lewison J in the case of *Thompson*. In that case Lewison J stated that

> the question is whether [the] occupation would have been obvious on a reasonably careful inspection. This is a relevant hypothetical question. Unlike paragraph 2(b) which requires inquiries to have been actually made, paragraph 2(c) does not require an actual inspection. It asks what would have been obvious if an inspection had been made.\(^\text{32}\)

Although this part of the judgment was obiter, its consistency with the judgment in *Thomas* is noteworthy. Commenting on the interpretation by Lewison J, Bogusz states that it focuses on only one aspect of the function of paragraph 2(c)(i), namely as an ex post legal test.\(^\text{33}\) If, she says, one reflects on the purpose of paragraph 2(c)(i) specifically, and on paragraph 2 of Schedule 3 as a whole, ‘there is a clear indication that it has a dual operational function’, one being the legal test and the other a point of reference for intending disponees as regards enquiries and the inspection process. Subsequently, when discussing *Thomas* and paragraph 2(c)(i), she states that the court

> has the benefit of hindsight when determining whether occupation is

\(^{32}\) Para 132.

\(^{33}\) Bogusz, *supra*, 280.
obvious, or occupation which may be considered obvious but which at the time could not be readily determined. The court, ex post and in full possession of the factual and contextual evidence can more readily make this judgment; by contrast to a surveyor or disponee who is piecing together the factual evidence and all the other relevant circumstances to determine if someone is in occupation. This leaves greater scope for error indicating that “obviousness” is a problematic concept.\(^3^4\)

Whilst Bogusz raises an important point, it is difficult to see why obviousness is necessarily a problematic concept. It must be borne in mind, first of all, that Ramsey J did not make a final determination, as Bogusz seems to suggest,\(^3^5\) but rather decided that Ms Thomas had ‘reasonable prospects’ of establishing that her occupation would have been obvious. Thus the ultimate outcome of the case at trial would be dependent on (among other things) the court accepting Ms Thomas’s submission that the builders and interior designers were working in the property both on her behalf and on behalf of Mr Burtenshaw.

However, if Ms Thomas’s submission were accepted then the question under paragraph 2(c)(i) would be whether the presence of the builders and interior designer, coupled with her almost daily visits, was such that it would have rendered her occupation obvious on a reasonably careful inspection of the property, taking into account the visible signs of occupation. If the physical presence of the builders and interior designers was found, in accordance with \textit{Lloyds Bank v Rosset}, to constitute presence on behalf of Ms Thomas as employer,\(^3^6\) there would be strength in the suggestion that her occupation, through her workers, was obvious; and the case for such a finding would be stronger taking into account her own presence at the property.

\(^{3^4}\) \textit{Ibid}, 282

\(^{3^5}\) \textit{Ibid}. She states that ‘As far as the court was concerned in \textit{Thomas}, Ms Thomas had a sufficient degree of occupation which would have been obvious on a reasonably careful inspection, by virtue of the builders and the interior design company involved the renovations of the property, and her regular visits to the property.’

\(^{3^6}\) See \textit{Thomas}, paras 22-23.
Under paragraph 2(c)(i) the court is required to put itself in the shoes of a disponee carrying out a reasonable careful inspection of the land at the relevant time and to come to a conclusion as to whether the occupation would have been obvious. This is, of course, an ex post test to be applied by the court; but it also provides a clear point of reference for disponees, and in light of Thomas and Thompson, it is submitted that the suggestion that there is no absolute duty placed on disponees to actually carry out a physical inspection is rather forceful. Certainly the express wording of paragraph 2(c)(i) does not require it. That having been said, it is clear that disponees fail to carry out an inspection at their peril, for otherwise they will inevitably run the risk of failing to identify those in occupation and thus miss out on the opportunity to enquire as to any interests which they may have.

The essential reason then why the ex post approach of the court and the concept of obviousness is not particularly problematic or disadvantageous for disponees is that the question of whether occupation would have been obvious will be looked at from the perspective of the disponee at the relevant time. No additional factors are required to be taken into consideration, and matters such as the intention and wishes of an alleged occupier are irrelevant, as the focus is simply on the visible signs of occupation. Hence Bogusz’s concerns about ‘occupation which may be considered obvious but which at the time could not be readily determined’ are misplaced. The upshot for disponees is simply that (whilst not compulsory) it would be prudent for them to carry out a reasonably careful physical inspection of the property. If they do so and the visible signs of occupation by a particular person are not obvious – meaning that there is no requirement to make an enquiry – then the relevant interest will not override the subsequent disposition. If they do not do so, they run the risk of the interest overriding in the event that occupation would have been obvious.

In the Thomas case, in the event of a final decision being made in favour of Ms Thomas, the reasoning of the court would thus be apt to be along the lines that the physical presence by the builders and/or the interior designers and/or Ms
Thomas sufficed to trigger the requirement to make enquiries, either of Ms Thomas herself or of the workers present on her behalf, on the basis that her occupation (whether by herself or through the workers) was obvious. However, whilst Ms Thomas may have had reasonable prospects of establishing that her occupation was obvious, there remains doubt over whether she would in fact have succeeded on this point. This is because of the possible argument that the physical presence of an agent or employee is (without more) insufficient to render the employer’s (i.e. Ms Thomas’s) occupation obvious.

Of course, a prudent lender may make enquiries of the workers, having noticed their presence, and such enquiries may lead them to find out about those in occupation and/or with interests in the property; equally, however, they may not, and for that reason it would arguably be wrong if the court were to find that, through the presence of her workers, Ms Thomas’s occupation was obvious. There is no requirement for a disponee to piece together ‘the factual evidence and all the other relevant circumstances to determine if someone is in occupation’; the occupation must be obvious based on the visual signs of occupation. On this view, the statutory exception and the message it sends to disponees is clear: the prudent disponee who carries out a reasonably careful physical inspection of the land, paying attention exclusively to visible signs of occupation, will experience few difficulties. By contrast, a disponee who does not carry out a reasonably careful inspection runs the risk of not identifying the visible signs of occupation that would trigger the need for further enquiry about the interests of the person concerned.

**Concluding remarks**

The judgment in *Thomas* (and also *Thompson*) provides a helpful and significant indication of how the courts will, and indeed ought, to interpret paragraph 2(c)(i) of Schedule 3 to the LRA 2002. As noted, although paragraph 2(c)(i) does not necessarily require an actual physical inspection of the land, it would undoubtedly be prudent for disponees to complete one. This is because in determining whether a person’s occupation would have been obvious on a
reasonably careful inspection of the land, the sole focus of the court will be on
the visible signs of occupation by that person. Thus because the court will
make its determination from a disponee’s perspective at the relevant time, the
sole focus of a disponee carrying out an inspection should also be on the visible
signs of occupation.

It follows from this that the court, in its ex post determination under
paragraph 2(c)(i), is not required to consider matters which would require
enquiry to ascertain them, such as the intention and wishes of the relevant
person, or other contextual evidence which may be relevant to the
determination of whether a person is in actual occupation. Nor, therefore, are
disponees required to consider such matters or to make other enquiries.

In reaching these conclusions, it is important to bear in mind the purpose of
paragraph 2(c)(i), namely to protect disponees where the fact of occupation is
neither subjectively known to them nor readily ascertainable. To be readily
ascertainable, occupation must, under paragraph 2(c)(i), have been visibly
obvious. If disponees were required to consider matters which require enquiry
to ascertain them, the purpose of paragraph 2(c)(i) would be undermined.
Accordingly, it is submitted that the law is clear and ought not to present many
difficulties – either for disponees in practice or for the courts in their ex post
application of it. The prudent disponee carrying out a thorough physical
inspection of the land at the relevant time, with a sole focus on the visible signs
of occupation, should experience few difficulties.

MATTHEW ASTLEY
There have been many debates concerning the management of company directors, particularly with regard to their corporate interests, terms of service and fiduciary duties. In order to explore these debates, it is necessary to investigate the general role of company directors and how they are regulated.

The problems that are associated with their fiduciary duties, terms of service and remuneration will also be examined along with the effect this presents on the general accountability of company directors. The solutions that have been suggested to resolve the inadequacies will also be evaluated in order to determine whether the existing system of controls and the vast quantities of recent corporate codes of practice adequately deal with the management of company directors.

**General role and regulation of Company Directors**

Company directors are typically appointed by shareholders to provide and maintain effective management of the company to which the shareholders are members.\(^1\) Indeed, in *Howard Smith v. Ampol*,\(^2\) Lord Wilberforce observed that directors may make decisions which they believe to be in the best interests of their company.\(^3\)

The issue of creating and maintaining good corporate governance is a feature that is common throughout the world and has been recently reinforced through numerous examples of major corporate collapses which include the failures of the Enron Corporation and WorldCom in America;\(^4\) the collapse of Parmalat in Italy; and the winding up of Maxwell Communications Corporation, Bank of Credit and Commerce International (BCCI), and Pretty Polly in the United

---

\(^1\) *Alexander Ward & Co Ltd v. Samyang Navigation Co Ltd* [1975] 1 WLR 673; *Shaw & Sons (Salford) Ltd v. Shaw* [1935] 2 KB 113 at 134.


\(^3\)*Ibid.* at 837; *Re Smith & Fawcett* [1942] Ch 304; s. 170(1) Companies Act 2006.

\(^4\) Another included the collapse of Tyco International Ltd.
Kingdom (UK) in the 1990s. All of which created an overwhelming impact upon their shareholders.

To prevent the reoccurrence of such events, there have been numerous provisions in different legal jurisdictions aimed at facilitating good corporate governance in order to adjudicate the various practices of corporate entities. Indeed, company directors are regulated in different ways in different legal jurisdictions: in America, directors are regulated by strict legislation under the Sarbanes-Oxley Act 2002, whilst other jurisdictions, such as the UK, operate largely through ethical practices as well as common law and statutory principles.

Indeed, it is the flexibility of such practices that has introduced numerous criticisms as to the UK’s progress in fostering the legitimate operation of corporate entities and the accountability of company directors’ towards their shareholders who are ultimately responsible for their appointment.

**Conflicts of Interest and Fiduciary Principles**

In the UK, whilst company directors are allowed to act in their interests to pursue the success of their company, case law has indicated that such interests may be exercised even where they are unpopular with the majority of the company shareholders. Indeed, in *Gramophone & Typewriter v. Stanley*, Buckley LJ acknowledged that company directors are not classed as servants or

---


6 The UK currently has 31 codes of practice, more than any other country: http://www.ecgi.org/codes/all_codes.php accessed on 8th April 2011.

7 Companies Act 2006.


9 s. 172 Companies Act 2006.


11 *Gramophone & Typewriter Ltd v. Stanley* [1908] 2 KB 89.
agents acting on behalf of their shareholders,\textsuperscript{12} whilst Cohen LJ in \textit{Grundt v. Great Boulder Gold Mines}\textsuperscript{13} observed that there is nothing unusual in refusing to permit shareholder interference in matters under the control of corporate directors.\textsuperscript{14} Indeed, Berle and Means\textsuperscript{15} observed that the ownership of widely dispersed shares among large American companies provided no individual or group control over the corporation which was instead administered by the corporation’s management.\textsuperscript{16} Consequently, it can be argued that an element of unfairness exists where shareholders wish to provide input as to how their company should operate since their interest may come into conflict with those of the directors who are regarded as providing the professional management of the company on behalf of the shareholders.\textsuperscript{17}

Gilson\textsuperscript{18} has commented that in the absence of shareholder constraint, it would not be unreasonable to expect company directors to abuse their power through maximising their own welfare as opposed to that of their shareholders,\textsuperscript{19} whilst Smith\textsuperscript{20} has observed that elements of greed exist among company directors since negligence and profusion always prevail when managing the finances of others.\textsuperscript{21}

Lipton and Rosenblum\textsuperscript{22} however, acknowledges that such arguments are unfounded since the occurrence of director misappropriation is rare\textsuperscript{23} even though they observe that, in strict legal terms, it should be the shareholders who

\textsuperscript{12} \textit{Ibid.} at 105-106.
\textsuperscript{13} [1948] Ch 145.
\textsuperscript{14} \textit{Ibid.} at 157; This principle also operates in Australia: \textit{Ashburton Oil NL v. Alpha Minerals NL} (1971) 123 CLR 614, Barwick CJ, at 620: ‘…Directors who are minded to do something which in their honest view is for the benefit of the company are not to be restrained because…shareholders holding a majority of shares in the company do not want the directors so to act.’
\textsuperscript{15} Berle, A, and Means, G \textit{The Modern Corporation and Private Property} (Macmillan Co, 1932).
\textsuperscript{16} They make this analogous to that of a captain and officers of a ship: \textit{Ibid.} at 135.
\textsuperscript{17} \textit{Id}; \textit{Gramophone & Typewriter Ltd v. Stanley} [1908] 2 KB 89.
\textsuperscript{19} \textit{Ibid.} at 836.
\textsuperscript{21} \textit{Ibid.} at Volume Two, pages 264-265.
\textsuperscript{23} \textit{Ibid.} at page 195.
ought to be regarded as the owners of the company in order to properly hold their directors to account in the event of any wrongdoing.\textsuperscript{24} Despite this, even though the law acknowledges that directors are required to exercise such interests in good faith so as to avoid any conflict,\textsuperscript{25} directors’ owe a fiduciary duty to act in the interests of their company since its undertaking is different from the entirety of the shareholdings.\textsuperscript{26} Consequently, it can be argued that although the law facilitates the proper management of directors' conflicts of interest it does not prevent directors’ from taking advantage of situations which they claim to be in the interests of the company.

Stokes\textsuperscript{27} has observed that the fiduciary principle under the law appears to show corporate directors acting in a position that is analogous to the situation whereby individuals are appointed to hold property on trust for others,\textsuperscript{28} however, directors do not hold legal title to their company property and as a result it can be argued that directors do not act as true trustees.\textsuperscript{29} Indeed, Bowen LJ in \textit{Imperial Hydropathic v. Hampson}\textsuperscript{30} observed that classifying company directors as trustees is not exhaustive of their position but is only useful for the purpose of observing the principles governing their actions.\textsuperscript{31}

It therefore appears that although company directors’ are not treated as trustees acting on behalf of their beneficiaries in the form of their shareholders, directors have a duty to act in the best interests of their company through personal

\begin{footnotesize}
\begin{enumerate}
\item \textit{Ibid.} at page 195-197.
\item Particularly with regard to their personal interests and that of the company: \textit{Industrial Development Consultants Ltd v. Cooley} [1972] 1 WLR 443 at 451; \textit{Guinness Plc v. Saunders} [1990] 2 AC 663; \textit{Kingsley IT Consulting Ltd v. McIntosh} [2006] EWHC 1288 (Ch); ss. 175-177 Companies Act 2006.
\item \textit{Ibid.} at page 161; These are the beneficiaries under a trust: \textit{Bristol & West Building Society v. Mothew} [1998] Ch 1, Millett LJ, at 1B: ‘...A fiduciary is someone who has undertaken to act for or on behalf of another in a particular matter in circumstances which give rise to a relationship of trust and confidence. The distinguishing obligation of a fiduciary is...loyalty...’; \textit{Hardoon v. Bellios} [1901] AC 118 at 123 ; s. 1 Recognition of Trusts Act 1987.
\item \textit{Re Exchange Banking Co (Flitcroft’s Case)} (1882) LR 21 Ch D 519 at 534; \textit{JJ Harrison (Properties) Ltd v. Harrison} [2001] EWCA Civ 1467.
\item \textit{Imperial Hydropathic Hotel Co Blackpool v. Hampson} (1883) LR 23 Ch D 1.
\item \textit{Ibid.} at 12.
\end{enumerate}
\end{footnotesize}
responsibilities that are certain and unavoidable regardless of the popularity of their actions which may not be supported by their shareholders. Consequently, it appears that the duties of company directors serve to exacerbate the problem regarding their terms of service towards their shareholders.

**Directors’ Terms of Service**

As shareholders contribute towards the capital of their company, it is considered to be an important factor in facilitating effective corporate governance to hold regular meetings between the company members and its directors to ensure that the latter is accountable to the former. Indeed, the 2003 Higgs report acknowledged that such meetings enable shareholders to develop a better understanding of any issues of the company from the perspective of its management while the Combined Code on Corporate Governance acknowledges that such meetings should be used to encourage the participation of investors with regard to the affairs of their company. Consequently, company directors in the UK are practically controlled and regulated through the general meetings with their shareholders.

The meetings between the company members and their directors are compulsory for companies with shares that are widely held, since the Companies Act 2006 require such meetings to be held every year with copies of the company’s annual accounts and reports being required to be disclosed by the

---

34 Higgs, D Review of the role and effectiveness of non-executive directors (Department of Trade and Industry, 2003), at page 67, para. 15.4.
37 s. 423 Companies Act 2006.
directors in order to review the business of the company\textsuperscript{38} as well the remuneration of its directors.\textsuperscript{39}

Despite this, it has been observed that in practice such meetings act as a limited form of control over the directors\textsuperscript{40} especially as shareholders are likely to trade-in their shares where they are despondent with the operation of their company,\textsuperscript{41} rather than voicing their point of view through their votes.\textsuperscript{42} Indeed, Esen has observed that part of the problem is caused by the shareholders themselves through devoting a small amount of time to attend the meetings, seek information regarding those meetings, and to participate in making company policies.\textsuperscript{43}

It can therefore be argued that the low attendance of shareholders at shareholder meetings serve to exacerbate the problem of controlling their directors,\textsuperscript{44} as well as illustrating the difficulty of current UK corporate governance in enabling shareholders to challenge the decisions of their directors. Indeed, Stokes has observed that the inadequacy appears to be based on the disparity between the legal assertion that shareholders control their company directors and the practical reality that shareholders fail to exercise their responsibilities of ownership.\textsuperscript{45} She therefore acknowledges that this inadequacy help to encourage corporate directors to act purely in their own interests rather than the interests of their shareholders who are merely reduced to acting as passive property owners.\textsuperscript{46} As a result, she opines that the power conferred upon

\textsuperscript{38} Ibid. at s. 417.
\textsuperscript{39} Ibid. at ss. 420-422; Directors’ Remuneration Report Regulations 2002 (SI 2002/1986).
\textsuperscript{42} Id.
\textsuperscript{46} Id.
company directors’ is likely to be unchecked and therefore illegitimate within the corporate framework.\textsuperscript{47} Indeed, in Re Smith & Fawcett,\textsuperscript{48} it was observed that there is an element of unwillingness by the judiciary in mediating commercial decision-making since directors are bound to exercise their powers in promoting the best interests of their company through their own considerations.\textsuperscript{49} Consequently, it can be argued that the problems associated with the accountability of company directors seem to be exacerbated on the part of company shareholders themselves based on their low attendance at shareholder meetings. As a result, this raises concern as to the legitimacy of the rate and level of directors’ remuneration since directors are required to disclose this information at shareholder meetings.

\textit{Directors’ Remuneration}

The subject of directors’ remuneration is another problem that is perceived to harm the accountability of company directors generally, particularly with regard to their rate and level of pay.\textsuperscript{50} Indeed, despite being aimed at ensuring effective corporate governance in the UK through the implementation of non-executive directors (NXDs), the Higgs report observed that corporate governance discussions tended to be narrowly focused around executive pay rather than those who ‘drive corporate success’.\textsuperscript{51}

Legally, it has been observed that company directors are not entitled to receive remuneration as of right since they are regarded as doing business for the benefit of their company.\textsuperscript{52} Indeed, in \textit{Hutton v. West Cork Railway},\textsuperscript{53} Bowen LJ

\begin{itemize}
\item \textsuperscript{47} \textit{Ibid.} at page 159; This may be a valid reason for the major corporate failures mentioned earlier.
\item \textsuperscript{48} [1942] Ch 304; \textit{Foss v. Harbottle} (1843) 67 ER 189 at 203; \textit{Burland v. Earle} [1902] AC 83 at 93-94.
\item \textsuperscript{49} \textit{Re Smith & Fawcett} [1942] Ch 304 at 306.
\item \textsuperscript{50} Parkinson, \textit{Corporate Power and Responsibility: Issues in the Theory of Company Law} (Clarendon Press, 1995), at page 236; Department for Business Innovation and Skills, \textit{A Long-term Focus for Corporate Britain: A Call for Evidence} (Department for Business Innovation and Skills, 2010); Department for Business Innovation and Skills, \textit{A Long-term Focus for Corporate Britain: Summary Responses} (Department for Business Innovation and Skills, 2011).
\item \textsuperscript{51} Higgs, \textit{D Review of the role and effectiveness of non-executive directors} (Department of Trade and Industry, 2003), at page 12, para. 1.9.
\item \textsuperscript{52} \textit{Re George Newman & Co} [1895] 1 Ch 674.
\item \textsuperscript{53} \textit{Hutton v. West Cork Railway Co} (1883) LR 23 Ch D 654.
\end{itemize}
commented that the mere fact that a person acts as a director does not imply an automatic right to be paid for their work.\textsuperscript{54} In practice however, the articles of association of a company will usually make provision for directors to receive remuneration.\textsuperscript{55} Indeed, a recent report published by the Department for Business Innovation and Skills (BIS)\textsuperscript{56} observed that the average remuneration of Chief Executive Officers (CEO) in FTSE100 companies increased by 13.6\% per year between 1999 and 2009\textsuperscript{57} resulting with a multiple of over 100 times the average employee earnings since 2002.\textsuperscript{58} It can therefore be argued that the rate and level of directors’ pay harms their general accountability.\textsuperscript{59}

To tackle this problem, the Combined Code on Corporate Governance\textsuperscript{60} has acknowledged that companies should avoid paying more to its directors than is necessary, in order to link individual performance with corporate success.\textsuperscript{61} The Code also commented that company directors should not be involved in determining their own level of remuneration.\textsuperscript{62}

To prevent directors’ input, the Code suggests that there should be a formal and transparent procedure whereby directors’ remuneration should be determined by a remuneration committee.\textsuperscript{63} Despite this, in its consultation paper, BIS found that although such committees now practically operate in numerous companies, they are not sufficiently independent or sensitive to the long-term interests of their companies and shareholders.\textsuperscript{64} Instead, BIS observed that the main driver behind the rate of pay is linked to the size of corporate entities since there is a

\textsuperscript{54} Ibid. at 672.
\textsuperscript{55} Companies Act 2006, Explanatory Note at para. 6 (316).
\textsuperscript{56} Department for Business Innovation and Skills, A Long-term Focus for Corporate Britain: A Call for Evidence (Department for Business Innovation and Skills, 2010).
\textsuperscript{57} Ibid. at page 26, para. 5.5.
\textsuperscript{58} Ibid. at page 25, para. 5.4.
\textsuperscript{60} Implementing the earlier Greenbury report on improving the remuneration system: Greenbury, R Directors’ Remuneration (Gee and Co Ltd, 1995).
\textsuperscript{62} Ibid. at page 14, para. B. 2.
\textsuperscript{63} Id.
\textsuperscript{64} Department for Business Innovation and Skills, A Long-term Focus for Corporate Britain: A Call for Evidence (Department for Business Innovation and Skills, 2010), at page 28, para. 5.11.
lack of evidence linking company performance to remuneration. Consequently, BIS commented that the lack of evidence prevents the long-term sustainable governance of corporate directorships.

Indeed, in March 2011, BIS found, from its responses to its consultation paper, that although there existed support for extending the membership of remuneration committees, a majority of non-supporters stressed that the implementation of such committees would have no impact on the levels of remuneration, as well as recommending that it should be the choice of individual companies as to whether or not to implement such committees where it is considered to be for the company’s advantage. It can therefore be argued that the current controls as to the level and rate of directors’ remuneration is likely to be a continuing problem in the future.

**Suggestions for Reform**

J.E. Parkinson has acknowledged that the general meetings of companies holding wide varieties of shares offer a very limited form of regulation over the proper management of directors’ interests which puts into disrepute their level and rate of pay as well as undermining their fiduciary duties. To resolve this, Parkinson proposes the existence of an independent organisation that is able to adjudicate directors’ terms of service, as well as their self-interested transactions. Parkinson further suggests that the organisation should also be empowered to enforce the fiduciary principle that is conferred upon corporate directors. Consequently, Parkinson opines that the existence of an independent body is likely to provide a satisfactory solution to the problem as well as

---

65 Ibid. at page 26, para. 57.
66 Id.
67 Department for Business Innovation and Skills, *A Long-term Focus for Corporate Britain: Summary Responses* (Department for Business Innovation and Skills, 2011), at page 21, para. 71.
68 Id.
70 Ibid. at page 236.
71 Id.
72 Id.
73 Id; s. 170(1) Companies Act 2006.
74 Id.
providing an appropriate mechanism to generally administer directors’ fiduciary powers.\textsuperscript{75}

In pursuance of this suggestion, numerous voluntary codes of practice have been published in the UK with the goal of fostering the legitimate operation of corporate entities for the benefit of their shareholders.\textsuperscript{76} Indeed, the Committee on the Financial Aspects of Corporate Governance\textsuperscript{77} commented that its Code aimed to improve the integrity and openness of corporate entities under the leadership of directors\textsuperscript{78} through the implementation of a minimum of three ‘special’ independent NXDs on the board.\textsuperscript{79}

The Code’s suggestion was later reviewed and accepted by Derek Higgs in a 2003 report\textsuperscript{80} but added that the independence of NXDs, and therefore the board generally, should be based on their objectivity in character, judgement and circumstances which could affect board’s management.\textsuperscript{81} This suggestion was later supported by the Financial Reporting Council (FRC) in its Combined Code,\textsuperscript{82} with the further recommendation that NXDs should be prepared to engage in communicative dialogues with major shareholders in order to ensure their interests are acknowledged by the board.\textsuperscript{83}

In addition to these recommendations, Lord Davies has recently suggested that the diversity of corporate boards could be further improved through the inclusion of more women who are likely to bring new ideas resulting in better

\textsuperscript{75} Id.
\textsuperscript{77} chaired by Adrian Cadbury.
\textsuperscript{78} Cadbury, A Report of the Committee on the Financial Aspects of Corporate Governance (Gee and Co Ltd, 1992), at para. 3.2.
\textsuperscript{79} Ibid. at para. 4.10.
\textsuperscript{80} Higgs, D Review of the role and effectiveness of non-executive directors (Department of Trade and Industry, 2003), at page 12, para. 1.14.
\textsuperscript{81} Ibid. at pages 36-37.
\textsuperscript{83} Ibid. at para. D.1.1, page 18; Higgs, D Review of the role and effectiveness of non-executive directors (Department of Trade and Industry, 2003), at page 31, para. 7.5.
decision-making. It therefore appears that the suggestions for reform aim, not only to improve the accountability of corporate directors, but also to enable corporations to become more socially responsible for their actions. Indeed, Buckley LJ in Re Horsley & Weight Ltd commented that the objects of a company can be philanthropic or charitable and need not be commercial provided that its actions are legal.

The John Lewis Partnership observed that being a good corporate citizen is linked to its commercial success by treating all of its partners with honesty, respect and fairness. The company also has a written constitution as well as a website that is devoted to maintaining its corporate governance in the long term interests of its stakeholders as well as managing its progress and performance. This partnership therefore illustrates the practical importance of maintaining good corporate governance in the UK.

Despite this, Stokes has commented that such practices do not reflect the corporatist vision of businesses whereby directors’ are empowered to account for a wide range of different matters rather than solely accounting for the interests of their shareholders. Indeed, Friedman has commented that the only social responsibility of corporations is to maximise their profits in order to create free and open competition. Indeed, it may be argued that this is reflected under s. 170(1) Companies Act 2006 which observes that the general duties of

---

84 Davies, M Women on Boards (Department for Business Innovation and Skills, February 2011), at pages 7-10.
85 [1982] Ch 442.
86 Ibid at 450.
87 http://www.johnlewispartnership.co.uk/Display.aspx?&MasterId=c489c8d5-be4b-483a-a3fb-65d3ab5688b1&NavigationId=1254 accessed on 18th April 2011.
89 http://www.johnlewispartnership.co.uk/Display.aspx?&MasterId=9066b519-cc0c-40ac-b641-5194e074c846&NavigationId=604 accessed on 18th April 2011.
90 Id.
92 Friedman, F. 'The Social Responsibility of business is to Increase its Profits' in Burchell, J (ed). The Corporate Social Responsibility Reader (Routledge, 2008), at pages 84-89.
93 Friedman describes the doctrine of corporate social responsibility as being 'fundamentally subversive': Ibid, at page 89.
directors are owed to their company as opposed to their shareholders, thereby illustrating the statutory practice of the corporatist vision of the company having priority over its shareholders’ interests. It may therefore be argued that UK law merely facilitates the interests of corporate directors rather than the interests of its shareholders in contrast to other countries such as France, Germany and the Netherlands where directors are required to be accountable to their investors.

It can also be argued that the various codes are not effective enough in tackling director accountability. Du Plessis has acknowledged that the ‘comply or explain’ nature of the codes may lead to numerous major corporate collapses in the future based on the lack of close scrutiny of their conduct by senior regulators. Indeed, Arora has observed that this observation coupled with the banking industry’s failure to observe the standards of good governance have greatly contributed to the current credit crisis.

The lack of mandatory provisions required to adhere to the various practice codes has also been acknowledged judicially. Indeed, Jonathan Parker J in Re Astec (BSR) Plc commented that the codes merely adjudicate the avoidance of unfairly prejudicial interests which may or may not exist in corporate matters. Although it can be argued that the codes are flexible enough to adapt to the required needs of time as illustrated by the recent Walker Review which aims to

---

94 s. 170(1) Companies Act 2006.
96 German Corporate Governance Code 2002 (as amended); Wooldridge, F and Davies, L. The German Corporate Governance Code’ (2010) (8) JBL 723.
97 Committee on Corporate Governance, Recommendations on Corporate Governance in the Netherlands: Forty Recommendations (25th June 1997), at para. 1.1.
100 Ibid. at 48.
102 Ibid. at 18.
104 Ibid. at 88.
reduce the risk of disastrous outcomes taken by bad strategic decisions in response to the recent banking crisis.\(^{105}\)

It therefore appears that effective UK corporate governance is emphasised through principles of self-determination as opposed to the imposition of statutory provisions, although from this, it can be argued that the fact further codes have been published which have merely repeated previous recommendations for reform, as illustrated by Cadbury, Higgs and the FRC above, suggests that the pace of ensuring effective corporate governance is slow,\(^{106}\) thereby requiring statutory regulation with fixed principles to deal with the conduct of corporations.\(^{107}\)

**The Companies Act 2006**

The main statutory provisions which seek to adjudicate the effective operation of corporate entities in the United Kingdom (UK) are contained in the Companies Act 2006.\(^{108}\) The Act was passed with the intention of reforming and updating company law to reflect modern needs as well as introducing greater flexibility as to the operation of corporate entities.\(^{109}\) Indeed, the Explanatory Note accompanying the Act acknowledges that as the business environment evolves, the risk of the legal framework becoming divorced from the needs of companies are likely to create obstacles in lawfully scrutinising the operation of corporations.\(^{110}\) Despite this, it has been observed that since the name of the Act has not changed from the previous 1985 Act,\(^{111}\) it might reasonably be implied

\(^{105}\) Walker, D *A review of corporate governance in UK banks and other financial industry entities: Final recommendations* (HM Treasury, 2010), at page 72, para. 5.13.

\(^{106}\) The Combined Code, published by the FRC, was initially published in 2000 and was later amended in 2003 and 2006. It was also revised in 2008: http://www.ecgi.org/codes/all_codes.php accessed on 18th April 2011.

\(^{107}\) Lee, P *Serving two masters - the dual loyalties of the nominee director in corporate groups* (5) (2003) JBL 449 at 468-469.

\(^{108}\) Companies Act 2006; The Act comprises 47 Parts, 1,300 sections and 16 Schedules, thereby making it the largest act ever enacted by Parliament: Omar, P. *In the wake of the Companies Act 2006: an assessment of the potential impact of reforms to company law* (2009) 20 (2) ICCLR 44.


\(^{110}\) Companies Act 2006, Explanatory Note para. 1(3).

\(^{111}\) Companies Act 1985; As a Bill, the Act began its parliamentary proceedings as the Company Law Reform Bill, although the name changed during the process in 2006: Legislative Comment,
that the intention of updating the law has been given a second priority in the sense of being classed as merely another Companies Act.\textsuperscript{112}

Despite this, the Act has tried to tackle the numerous problems of associated with the accountability of company directors. Indeed, the 2006 Act has, for the first time in company law history, codified some of the common law principles relating to directors’ duties\textsuperscript{113} which include, inter alia, a duty in promoting their company’s success\textsuperscript{114} and a duty to employ reasonable care, skill and diligence when executing their functions.\textsuperscript{115} This has the intended effect of making company directors more accountable for their actions, as well as being more exposed to an increased risk of litigation in the event of a failure to comply with the Act\textsuperscript{116} since shareholders are provided with an extended power to sue their directors for breach of duty, negligence and default in a derivative claim\textsuperscript{117} under s. 260(3).\textsuperscript{118}

Despite this, the Act acknowledges that the general duties of company directors are owed to the company and not directly to its shareholders.\textsuperscript{119} Indeed, s. 172 displays subjectivity with regard to directors’ interests since they are able to act in a way ‘they consider’ to be for the benefit of their company.\textsuperscript{120} Consequently, it can be argued that directors continue to enjoy extensive powers while

\textsuperscript{113} s. 170(3)-(4) Companies Act 2006; Omar, P. ‘In the wake of the Companies Act 2006: an assessment of the potential impact of reforms to company law’ 20 (2) (2009) ICCLR 44.
\textsuperscript{114} s. 172 Companies Act 2006.
\textsuperscript{115} Ibid. at s. 174.
\textsuperscript{116} For example, s. 183 Companies Act 2006 states that directors who fail to declare the extent of their interest that has been entered into by the company will be liable to receiving a financial conviction.
\textsuperscript{118} s. 260(3) Companies Act 2006.
\textsuperscript{119} Ibid. at s. 170(1).
\textsuperscript{120} Ibid. at s. 172(1).
shareholders may encounter a potential struggle to hold their directors’ to account.\footnote{121}{Birds, J et al. Boyle & Birds’ Company Law (7th ed. Jordan Publishing, 2009), at page 357, para. 11.1.}

Indeed, Lee\footnote{122}{Lee, J. ‘Shareholders’ derivative claims under the Companies Act 2006: market mechanism or asymmetric paternalism?’ (2007) 18 (11) ICCLR 378.} has observed that the Act has not produced a breakthrough in relation to directors’ duties since it requires the court to interpret the duties in accordance with common law and equitable principles thereby providing a lack of discretion to establish new duties under the Act since judicial precedents are required to be followed.\footnote{123}{Ibid. at page 391.} This analysis has also been qualified by Lord Hodge in Eastford v. Gillespie,\footnote{124}{Eastford Ltd v. Gillespie [2010] CSOH 132.} where he commented that there is nothing in the statutory provisions which has indicated that Parliament intended to alter the pre-existing rules with regard to interpretation.\footnote{125}{Ibid. at para. 13.} Despite this, Lord Hodge commented that although the courts are required to treat the general duties in the same way as the pre-existing rules, the interpretation of the statute will be able to evolve since the courts are also required to take into account the continued development of the non-statutory law.\footnote{126}{Id.}

It therefore appears that although the Act has made the general duties of directors more accessible to shareholders who will be able to make a derivative claim against their directors, it can be argued that the Act has not provided much of a difference with regard to formally acknowledging directors duties being owed directly to the shareholders themselves. Indeed, the Act illustrates that directors’ continue to owe their duty to the company and may exercise their interests in a manner that they consider to be for the benefit of their company. With regard to directors’ remuneration, the 2006 Act has reinstated the Directors’ Remuneration Report Regulations 2002 which was initially inserted into the Companies Act 1985\footnote{127}{Directors’ Remuneration Report Regulations 2002 (SI 2002/1986); ss. 7(A) and 257 Companies Act 1985 (repealed).} requiring companies to hold a compulsory
annual shareholders’ vote on directors’ remuneration packages.\(^\text{128}\) Indeed, Paragraph 82 of Table A under the Companies (Tables A to F) Regulations 1985 originally provided that it was ‘the company’ which determined the remuneration of directors through the passing of an ordinary resolution.\(^\text{129}\) The 2006 Act, however, has gone further since the Secretary of State has produced the Companies (Model Articles) Regulations 2008\(^\text{130}\) which now allows the directors themselves to determine their remuneration for companies that have been formed under the Act.\(^\text{131}\) Indeed, the 2006 Act also requires directors of publicly quoted companies to prepare a remuneration report for each financial year of the company’s existence under s. 420,\(^\text{132}\) whereby a failure to comply will result with the director receiving a financial conviction.\(^\text{133}\)

It has been observed that allowing directors to set their own rate and level of pay has provided a curious legal change especially with regard to the public concern in recent years.\(^\text{134}\) Indeed, it can also be argued that the 2008 Regulations has expressly acted against the Combined Code on Corporate Governance which stresses the requirement that no director should be involved in determining their remuneration as this would produce a more formal and transparent procedure in determining remuneration packages as well as ensuring effective corporate governance.\(^\text{135}\) This result therefore illustrates the voluntary nature of complying with the various codes of practice which can be lawfully disregarded. Indeed, Gompertz described the various codes as having a ‘light-touch’ to regulation,\(^\text{136}\) which is in comparison with other jurisdictions such as Germany whereby certain terms of the German Corporate Governance Code 2002 (as

\(^{128}\) Id.

\(^{129}\) Companies (Tables A to F) Regulations 1985 (SI 1985/805), Table A para. 82.

\(^{130}\) Companies (Model Articles) Regulations 2008 (SI 2008/3229).

\(^{131}\) Ibid. at Scheds.1 and 2, para.19, and Sch.3 para.23; Although earlier companies are free to adopt the model articles: Ibid. at Explanatory Note.

\(^{132}\) s. 420(1) Companies Act 2006.

\(^{133}\) Ibid. at s. 420(3).


\(^{136}\) Gompertz, K. ‘A little more power to the people?’ (2009) 14 (2) Cov L J 28
amended) contain provisions that companies are compelled to observe under relevant law.  

The European Commission has issued a Recommendation aimed at resolving the problems associated with directors’ remuneration within the European Union. Indeed, the Commission suggested that shareholders should be informed of the earnings of individual directors in order to ensure adequate control is exercised by shareholders over such matters. Despite this, the Recommendation is non-binding in UK law and it is this result that has led Smerdon to suggest that it has created little impact in UK corporate governance.

Consequently, it appears that rather than tackling the problems associated with directors’ remuneration, the 2006 Act promotes the continuance of the problem by allowing directors to set their own remuneration. Although it can be argued that such directors will be subject to the fiduciary principles of good faith with regard to promoting the success of the company for the benefit of its members under s. 172 of the Act, and will be therefore exposed to the risk of litigation by the shareholders.

Conclusion

In conclusion, it appears that the existing controls of company directors’ interests, duties and remuneration through the general meeting of company shareholders are inadequate since the structure of UK corporate governance is based on self-determination rather than mandatory provisions that are designed to combat the problems that are associated with directors’ interests.

137 German Corporate Governance Code 2002 (as amended); Wooldridge, F and Davies, L. 'The German Corporate Governance Code' (8) (2010) JBL 723 at 724.
139 Ibid. at para. 6.4.
140 Ibid. at para. (2).
Consequently, the self-determinative nature of the various codes of practice, serve to prolong the unaccountability of company directors towards their shareholders, thereby requiring more to be done statutorily to ensure director accountability.

Although, the Companies Act 2006 has tried to make a conscious effort to keep company law relevant to current needs, the Act has not produced great change in tackling the problems associated with directors’ accountability particularly with regard to the issue of remuneration, which illustrates the incompleteness of the statutory code.\textsuperscript{144} This is further illustrated by the plethora of recent codes of practice which voluntarily, rather than mandatorily, tries to balance directors’ managerial power with the interests of shareholders which creates a significant gap between the interdependency of the codes and the law.\textsuperscript{145} The lack of statutory force in implementing the codes therefore create uncertainty in UK corporate governance although only time will tell in determining how effective the 2006 Act will be at ensuring effective governance in the future.

ANDREW PELLINGTON

\textsuperscript{144} Omar, P. ‘In the wake of the Companies Act 2006: an assessment of the potential impact of reforms to company law’ 20 (2) (2009) ICLR 44.
Risky Business: Challenges and Changes in the Assessment of the Standard of Care Owed to Participants in Supervised ‘Extreme’ Sports

FREDERICK LYON

Introduction
It has proved very difficult to categorise extreme sports. It is a phrase used widely to indicate risk and elements of counter-culture. The sports (or more accurately activities) referred to generally, by their nature, involve an element of risk seeking. This risk taking is either the main reason for engaging in the activity, or at the least, forms part of the rationale for engaging in it. The problem with this definition is that it could include sports usually categorised as ‘mainstream’ such as rugby or boxing where an unavoidable element of risk persists that adds to the excitement of participation. The definition is therefore inherently inexact. This article will predominantly deal with six activities namely; mountaineering, outdoor rock climbing, canoeing, kayaking, and off-piste skiing. The feature held in common by these ‘extreme sports’ is that participants pit themselves not merely against others, but against risks posed by natural phenomena themselves. Consequently they are among those activities with a high number of uncontrollable variables, and inevitable risks that are challenging to calculate.

It is the uncontrollable and unpredictable nature of such activities that has provided, and continues to provide, the law with significant difficulties when assessing whether tortious wrongs have been committed. The law has attempted to steer a course between a paternalistic attitude to risk and allowing personal autonomy. There has been significant support for the latter position from the

1 Freelance outdoor instructor and member of the Mountain Leader Training Association, holding the Mountain Leader Award (Summer), Single Pitch Award, British Canoe Union Level 2 (Kayak and Canoe) and Mountain Bike Leader Award. BPTC student at City University.
2 They are perhaps sports only in the sense referred to by Barnaby Conrad (popularly misattributed to Ernest Hemingway) when he stated “there are only three sports: mountaineering, bullfighting, and motor racing; all the rest are merely games”.
3 ’Adventurous outdoor activities’ would perhaps be a more suitable definition.
media, and from government advisors. These efforts to avoid a so-called “compensation culture” have influenced judicial decision-making and statutory legislation. Cases involving the autonomous action of an unsupervised person are relatively settled (certainly in relation to occupier’s liability), if responsibility has not been taken for their welfare and any hidden dangers have been identified then there is very rarely any position in which a duty is owed. As Lord Hoffman stated in Tomlinson v. Congleton Borough Council “if people want to climb mountains, go hang-gliding, or swim or dive in ponds or lakes, that is their affair.” Where people take responsibility for an activity, putting themselves in the position of instructors, the position in law is rather more complex.

There are no clear figures for the number of injuries and deaths that occur in supervised adventurous activities each year. While it is likely that in percentage terms the number is very small indeed, each case is nonetheless of obvious significance to those who are unfortunate enough to be involved in one. It has long been held that in such cases a duty is owed by the supervisor: by taking responsibility for the welfare of a client, the instructor makes himself or herself liable for negligent acts or omissions which lead to loss and damage. It is in assessing the nature of this duty and the standard of care required that special considerations must be made. This article deals with two main difficulties in cases involving supervised extreme sports. The first is a point, most usually raised by a claimant, that an instructor is liable if their standard of care fell below that which would be expected of a reasonable instructor. It is submitted that to argue this way, placing total reliance on this modified Bolam test, is to oversimplify cases of this nature and could lead to the claimant being unable to recover. The second point, usually relied on by the defendant, is that by

---

5 Common Sense, Common Safety, Lord Young [2010].
7 The Compensation Act 2006.
consenting to participation the claimant consented to all risks and hazards associated with the activity. This is usually raised in order to demonstrate that a lower standard of care is owed to the claimant. It is submitted that a great deal of care must be taken with this approach. A myriad of factors, particularly prior experience, must be taken into account: an over-reliance on the consent afforded by participation in a high-risk activity, though it has had some success in the past, is a potentially poor strategy. It will be suggested that the reliance placed on the factors necessary to calculate a standard of care has shifted and continues to shift in line with public policy. To this extent, the standard of care required of instructors or guides in these activities also alters.

**CLAIMANT**

The *Bolam test*

Where the defendants are professional guides and instructors, a modified version of the *Bolam*\(^{12}\) test is an attractive prospect to a claimant. Essentially it will set the standard of care at the level of a reasonably careful and competent instructor undertaking the activity in question.\(^{13}\) In *Anderson v. Lyotier*\(^{14}\) a ski instructor was found to be negligent as he had allowed an inexperienced skier under his instruction to go down a steep off-piste run. Fosket J, while noting that the modified *Bolam* test was only of limited use, made a finding that the instructor had been negligent almost entirely on that basis. It is worth noting, however, that leave to appeal the case was subsequently granted but the case was compromised prior to being heard in the Court of Appeal.\(^{15}\) A difficulty that arises in such a case is that, unlike in medicine, where operating outside of accepted practice always foreseeably results in injury, the same is not true in activities with a high objective risk in any event. The fallacy is to assume that once a breach of standard practice is found, this is sufficient to amount to actionable negligence. Due to the high-risk nature of these activities the duty of an instructor can only be the duty to minimise the risk of injury or death, not to prevent it. It is nonetheless a useful starting point to assess a standard of care.

---

\(^{12}\) *Bolam v Friern Hospital Management Committee* [1957] 2 All ER 118.

\(^{13}\) Woodroffe-Hedley v. Cuthbertson (1997, Unreported), in a case involving mountaineering.

\(^{14}\) *Anderson v. Lyotier* [2008] EWHC 2790, Per Foskett J.

\(^{15}\) The Duty of Care Owed by Instructors in a Sporting Context, William Norris QC.
Assessment under the Bolam test

Regulation

In order to assess the negligence of an instructor there are reams of regulatory and licensing law governing the provision of adventurous activities. The system is currently overseen by the Adventurous Activities Licencing Authority (AALA).\(^{16}\) While failures to conduct formal risk assessments or to comply with necessary regulations can help a claimant’s case against an instructor or organisation, these are not enough on their own. Indeed, where risks are variable, as they so often are in the outdoor environment, they are only of limited use.\(^{17}\) In the case of Uren v. Corporate Leisure, a case involving an ‘it’s a knock out’ style game where contestants attempted to get objects in an inflatable filled with around 8” of water (although not a designed paddling pool) as quickly as possible, Atkins LJ noted that “there can be no legal definition of ‘acceptable risk’ in relation to a particular game. It must depend on the activity concerned, the way in which it was being undertaken, who was doing it and in what circumstances.”\(^{18}\) Any breaches of regulations can therefore only be indicative of negligence, even if they establish that an instructor or organisation did not act as a reasonably careful and competent one would have done. It is of note however that this argument seems to hold more weight in the context of activities with fewer variables: the very recent case of Pinchbeck v. Craggy Island, involving an fall of 5-6ft on to a mat at an indoor climbing wall, suggests that where an instructor takes responsibility for the welfare of a claimant their failure to comply with health and safety regulations and set procedures can be sufficient indication on which to base a case.

\(^{16}\) It should be noted however that while AALA have been permitted to continue providing this service they are being phased out in favour of a recognised, but voluntary, code of practice (See Lord Young report, Common Sense, Common Safety). This author anticipates that points raised in this article will become even more relevant with the decline in the number of regulations. It is also anticipated that feeling out the extent of the codes of practice may lead to further cases in this area been taken to trial.

\(^{17}\) Uren v. Corporate Leisure (UK) and another, [2011] EWCA Civ. 66, 2011 All ER (D) 49 (Feb), per Smith LJ at paragraphs [41]-[42].

\(^{18}\) Uren v. Corporate Leisure (UK) and another, [2011] EWCA Civ. 66, 2011 All ER (D) 49 (Feb), per Atkins LJ at paragraph [76].
Weighing up severity and likelihood

The first challenge that must be attempted by the court is to carry out a risk assessment that should have been taken by the instructor at the time of the accident. In doing this the court must weigh up both the severity and likelihood of the risk and balance this against other factors such as the cost and difficulty of prevention and whether preventing the risk would prevent the undertaking of a desirable activity.\textsuperscript{19} In cases involving outdoor adventurous activities it would seem that the first part of this assessment (and indeed often the only part considered by instructors making on the spot or written risk assessments) may be difficult to meet. The activities produce very serious risks: serious injury or death is a foreseeable possibility in any cases of rock fall, avalanche, or trapping underwater. The likelihood of the risk is also challenging. Many of the risks are inherently unpredictable. Establishing a likelihood of a risk occurring can take years of experience and careful consideration, and depends entirely on the conditions encountered by the instructor on the day. Finally, because they are inherent in the activity, the costs involved in risk prevention are incalculable: it may be possible to minimise a risk of head injury with a helmet but it is not possible to eliminate it completely if the client is hit with a big enough rock. Given this element of unpredictability and the level of risk involved, taking a risk-averse attitude would lead to the activity never being undertaken. Needless to say, this has not been the preferred stance of the courts.

Reasonableness

Using the formula above it is possible to begin to make a calculation of the reasonableness of the instructor acting as they did. It is also necessary to take into account the circumstances in which the accident occurred. As was noted in the case of Cordon v. Basi by Sir John Donaldson MR, “...you are under a duty to take all reasonable care taking account of the circumstances in which you are placed.”\textsuperscript{20} In that case it was shown that where a decision is made in the heat of the moment a lower standard of reasonableness applies. This has been applied

\textsuperscript{19}This is somewhat reminiscent of the formula posited by Learned-Hand J in the US case of United States v. Carroll Towing Co (1947) 139 F 2nd 169, that an acceptable risk is one when the costs of prevention are less than the severity of any foreseeable injury times its likelihood.

\textsuperscript{20}Cordon v. Basi [1983] 2 All ER 453 per Sir John Donaldson MR
numerous times in a sporting context, in horse riding\textsuperscript{21} and football,\textsuperscript{22} for example. It is reasonable to think that this lower standard of reasonableness would apply to many decisions made in the outdoor environment, and indeed it does. One need only need to read D Cotton QC’s judgment in \textit{Pope v. Cuthbertson} to see that where a case relied on split second reactions – in that case taking in a rope during a fall – that the test for what is reasonable in those circumstances is very high indeed.

It must, however, be recognised that many mistakes made by instructors are made long before their client is ever injured. Failing to read the weather conditions or taking an inexperienced client to the wrong venue would both be potential examples of negligent mistakes to which this reasoning would not apply. Even during the activity the courts seem to have been reluctant to find that an instructor acted in the heat of the moment. In another case involving a guide named Cuthbertson (seemingly not the same unfortunate man), the line between what was reasonable or not was vanishingly small. Mr Cuthbertson was engaged to take a client, Mr Woodroffe-Hedley, to the summit of the Tour Ronde, a mountain in the Mt Blanc Massif. The route they took, although on the north face and involving technical climbing, was not beyond Mr Woodroffe-Hedley’s ability and was a popular choice (there were two other parties climbing it that day). In terms of venue and planning before the climb neither side could fault Mr Cuthbertson. Client and mountaineer were roped together; they moved in pitches with the leader climbing the full rope length and setting up an anchor point before the second came up to join him. It was on the setting up of these anchor points that the case rested. As they neared the upper face, Mr Cuthbertson became concerned that the heat of the morning sun might dislodge rocks from the terrain above. In accordance he planned to move quickly. He arrived at a place to set an anchor point and used a single metal ‘screw’ (looking like a hollow tube) placed in the ice. It is industry practice to use two at least. All his others were blocked with ice, making them impossible to place. He stated that he did not want to spend time clearing a second and made a decision not to

place it. The court held that Mr C was negligent in this decision. The consequences of his actions were serious. As he moved off the anchor, having attached Mr Woodroffe-Hedley to the screw, a sheet of ice broke off beneath his feet – carrying him with it. The added pressure shattered the ice around the screw, and both Mr Cuthbertson and Mr Woodroffe-Hedley fell down the mountain. The latter was killed; the former sustained a severely broken leg.

The argument that Mr Cuthbertson made a split-second judgement based on the risk of rock fall and his experience as a climber, and that this judgement was made in the heat of the moment, was specifically rejected by Dyson J. He held that Mr Cuthbertson had time to think about his decision and made a negligent choice given the various risks involved.23

*The client*

Who the client is will be a consideration in all cases involving allegations of the negligence of outdoor instructors and guides. It is trite law to say that in practice the courts are more likely to impose a higher standard of care on instructors supervising children than on those supervising adults. It should also be noted that different standards of care will apply to different clients depending on their experience. While it would, for example be unacceptable to allow a novice to descend a grade 5 rapid (in a 1-6 scale of difficulty), the same could not be said for someone who the instructor was aware had built up a certain level of skill and was capable of doing so, provided they were supervised appropriately. It is not that the participant has consented to the risks confronted on the grade 5 rapid, but that the instructor has taken appropriate steps to minimise the risks of the activity while still allowing the client to take part at the correct level for them. The instructor by making an appropriate assessment of the capability of the client does not materially increase the risk of the already risky activity.

As can be seen, a modified *Bolam* test acts as a useful starting point, taking into account as it does the nature of the activity, who was taking part and in what

---

23 It is of note that the British Mountain Guides professional standards committee exonerated Cuthbertson and he was permitted to continue working as a guide: ‘Top Climbers Clear Scot of Blame in Fatal Fall’, The Scotsman, 7th October 1997.
circumstance. However, it cannot be wholly sufficient to demonstrate that an instructor was negligent. It must also be shown on the balance of probabilities that the decision not only failed to accord with best practice but also put the client at a foreseeable and probable increased risk of injury.\textsuperscript{24} In activities where the risks are objectively higher in any event, this may leave a claimant in difficulties.

**Material increase in risk**

As other authors have noted, the mere fact that a leader acts as no other reasonable leader would have done does not establish negligence unless they foresee or should have foreseen that a higher risk of injury was likely to follow from their acts or omissions. As the case of *Whippy v. Jones* clarified, the possibility of a risk materialising is not enough: there must be a sufficient probability to lead the reasonable person to anticipate that it would.\textsuperscript{25} As William Norris QC suggested, there must be close analysis of the extent of any duty and the foreseeability of harm. This must be carried out not simply when seeking to establish causation or look into remoteness of damage, but also to establish the standard of care that should have been afforded to the claimant.\textsuperscript{26}

In sports with a comparatively high number of objective risks, making such an assessment is both necessary and extremely tricky. This challenge is well illustrated by the case of *Chalk v. Ministry of Defence*, which represents the school of thought that goes beyond the *Bolam* test. In this case an inexperienced climber was being trained as a member of mountain rescue by the RAF. The leader of the party checked the weather and made a decision to climb on the Red Tarn Face of Helvellyn in the Lake District. Experts in the case confirmed that this was not the best choice given the conditions on the day,\textsuperscript{27} as snow build-up – the result of prevailing winds – had resulted in conditions that were primed for what is known as a ‘wind slab avalanche’.\textsuperscript{28} Having climbed a route on a rocky buttress the party were descending when they were avalanched. It was held that given

\textsuperscript{24} Climbing Accidents – the Duty and Standard of Care, Matthew White. The Duty of Care Owed by Instructors in a Sporting Context, William Norris QC.

\textsuperscript{25} 2009 EWCA Civ. 452, per Aitkins LJ.

\textsuperscript{26} Per Nelson J at paragraph 34.

\textsuperscript{27} Per Nelson J at paragraph 46.

\textsuperscript{28} Per Nelson J at paragraph 34.
the conditions, and in spite of the fact that climbing in that location was not the best choice in the circumstances, the injury was nothing more than an unfortunate accident. The instructor had taken all reasonable steps to minimise the risk short of not going to the cliff at all.29

**Expert Evidence**

Given the very challenging nature of the assessment that must be made, first of the actions of the reasonable instructor in all the circumstances and second of whether that action materially increased the risk involved, the importance of high quality expert evidence cannot be overstated. The cases in this area often seem to turn almost entirely on the opinion of experts in the field. Decisions by solicitors and counsel with regards to the selection of experts should therefore be carefully considered. Pity for example the claimant’s legal team in the case of *Pope v. Cuthbertson*.30 Though their expert knew the local area extremely well and had been climbing a long time, he was not an instructor, and in fact based a number of his submissions related to industry practice on a book written by the defendant’s expert. It is unsurprising that the judge found the defendant’s expert significantly more compelling.

**DEFENDANT**

**The issue of consent**

It is common in cases involving supervised extreme sports for the defendants to advance the case that a lower standard of care was owed to the claimant on the grounds that they had consented to the risks inherent in the activity by participating. As with the *Bolam* test it seems an intuitive and useful argument to make. It is, however, once again not as simple as it first appears. I submit that in fact the argument requires an assessment of a number of issues, and in any event only adds anything to a defence in a limited number of circumstances.

First, it is clear that cases involving supervised high risk activities and cases involving the same activities unsupervised require different arguments. Where

---

29 Chalk v. Ministry of Defence [2002] All ER (D) 223 (Mar), per Nelson J.
an activity is unsupervised the courts have tended, with good reason, to take the view that most adults have sufficient autonomy to consent to their own risks, particularly if the risk is obvious.\(^{31}\) A supervised activity is a different situation and different arguments must be advanced accordingly. This was acknowledged in the case of \textit{Trustees of the Portsmouth Youth Activities Committee (A Charity) v. Poppleton}, which concerned a young man climbing unsupervised on an indoor wall.\(^{32}\) An example of the alternative approach can be found in the case of \textit{Fowles v. Bedfordshire County Council} where a trampoline instructor was found liable having taken responsibility for the welfare of a student while teaching them a somersault.\(^{33}\) It should be noted however that in that case the claimant was a child and the courts were therefore more likely to impose a higher duty of care.

The argument for suggesting that people can consent to all risks not caused by the negligence of their instructors is clear. As May LJ stated in \textit{Portsmouth Youth Activities v. Poppleton}, “Adults who choose to engage in physical activities which obviously give rise to a degree of unavoidable risk may find that they have no means of recompense if the risk materialises so that they are injured.”\(^{34}\) This is echoed in many other cases involving instructors. In \textit{Woodroffe-Hedley v. Cuthberston}, Dyson J stated: “mountain climbing is extremely dangerous. That is one of the reasons why so many risk their lives every year on mountains. Anyone who climbs with a guide is as a matter of law consenting to the ordinary dangers of mountain climbing.”\(^{35}\) In \textit{Pope v. Cuthbertson}, D Cotton QC stated: “If you are going to engage in rock climbing you must acknowledge and accept the risks you take and not offload your responsibility to others.”\(^{36}\) The same view has also


\(^{32}\) Trustees of the Portsmouth Youth Activities Committee (A Charity) v. Poppleton [2008] EWCA Civ. 646, 2008 All ER (D) 150 (Jan), Per May LJ at paragraphs [8], [17] and [20]. It is of interest to note the as yet unreported case of Pinchbeck v. Craggy Island (2012) EWHC in this regard. Here a fall of significantly less severity than that experienced by Mr Poppleton was suffered by the claimant. It was held by HHJ Patrick Curran QC (sitting as a judge of the High Court) that Mrs Pinchbeck was only 1/3 negligent, the instructor having taken responsibility for her welfare and the mats providing a hidden danger in making her think it was safe to jump. This author will watch with interest to see if the case progresses to appeal.


\(^{34}\) Trustees of the Portsmouth Youth Activities Committee (A Charity) v. Poppleton [2008] EWCA Civ. 646, 2008 All ER (D) 150 (Jan), Per May LJ at paragraph [1].

\(^{35}\) Woodroffe-Hedley v. Cuthbertson (1997, Unreported) per Dyson J.

\(^{36}\) Pope v. Cuthbertson (1995, Unreported) per D Cotton QC.
been voiced by academics: “the injured party has undertaken an activity, which in its own way, is fraught with danger, and in which it is generally known that injury may be caused without negligence on the part of anyone.”

There is, I submit, a fly in this ointment used by defendants. The consent must be both full and informed to apply: in many such activities where a client is inexperienced, this ideal does not seem to be a possibility. It is industry practice to have clients sign a participation statement consenting to the general risks of the activity. The standard template for this is the British Mountaineering Council (BMC) participation statement, which is generally modified to suit the activity in question. The statement broadly indicates that the client acknowledges that the activity in question is dangerous and carries the potential consequence of injury or death. The BMC makes clear in its advice to instructors that it is important that participants in adventurous activities are made aware of this. It does not, however, go further than this. No mention is made of explaining the types of risk to the client or explaining the chances of the worst occurring. While in practice instructors and guides may give specific warnings, this is not a mandatory requirement. Furthermore, if the client is notified of the risk during the activity, then it will depend on the circumstances as to what extent they could fully consent. If informed of the risk of an avalanche while walking, when they could only reasonably be expected to rely on the expertise of this instructor, consent to such a risk could only be invalid. This is particularly true given the complex nature of such phenomena.

Many of the risks associated with adventurous activities are difficult to predict and can take years of experience to understand properly. They are most certainly not automatically obvious risks for the lay observer. There are many reasons why people employ guides and instructors, but perhaps foremost among them is that they don’t have the expertise to be aware of, read and manage these risks themselves. There are echoes of the case of Kirkham v. Chief Constable of

---

37 Torts: Commentary and Materials, 10th Edn. Sappenden et al. (paragraph 16.40).
38 How Liable are You?, The BMC, 2000.
Greater Manchester,\textsuperscript{39} involving the suicide of prisoner in a police cell, in that the act that the defendants rely upon in demonstrating that the claimant acted voluntarily is the very act that the defendant’s duty required him to prevent (or at least minimise the risk of). The point that the claimant consented is however still often mentioned by judges. The case of Pope v. Cuthbertson involved a client being taken rock climbing and being allowed to place their own protection that would halt them in the event of a fall.\textsuperscript{40} This was supervised. When the claimant did fall, one of the pieces of protection ripped out and he hit the ground. There were multiple issues in that case in relation to the experience of the claimant and whether Cuthbertson should have allowed him to lead at all. D Cotton QC on the one hand was adamant that the claimant had consented, while still remaining vague as to how the protection failed. He stated it was either due to inexperience or bad luck. If inexperience was the issue then it must be recognised that by allowing the claimant to place his own protection, Cuthbertson took the chance of materially increasing the risk. Ultimately, given the facts available in the judgment, the case was decided correctly due to the uninstructed experience the claimant was found to have, and the measures Cuthbertson took to test this experience before-hand. Mr Cuthbertson increased the risk, but did not act outside industry practice in the circumstances; furthermore, given the measures taken the resulting consequence was possible but by no means probable. The reasoning employed – that simply by acting as the claimant did in climbing, he consented – is nonetheless flawed.

In some cases it is clear that an experienced claimant does consent to risk, they are fully informed of the potential consequences of their actions and understand that they put themselves in the way of certain unavoidable but identifiable risks. There will also, of course, be some cases where the risk was so obvious the claimant can be taken to have consented to it simply by participating (walking over a cliff, as an intelligent adult, would be an example); but this is by no means always the case. Care should be taken by practitioners to find out the exact nature of the claimant’s experience, to assess what risks they were informed of,

\textsuperscript{39} [1990] 2 QB 285.
\textsuperscript{40} (1995, unreported).
and to find out the manner in which they were informed. In order to rely on consent to lower the standard of care a large amount of information needs to be taken from both claimant and defendant. Given the extra time, effort, and expense for the client required to establish any consent, it should be seriously considered whether it is worthwhile to run the argument at all. If the defendant is negligent, both materially increasing the risk and acting in a way no reasonably careful and competent instructor would have done then they are simply negligent and the claimant will not be able to consent to that negligence. If, on the other hand, they do not fall within the tests set out above then there is no actionable negligence in the first place. There is in many cases obviously mileage in a contributory negligence argument, but where the risk is one created by natural phenomena (as in Cuthbertson) then it is likely that no argument related to contributory negligence could be run in any event. In cases where consent is a difficult issue one must question in every case what it adds to the defence.

**Would the steps required to prevent a risk prevent a desirable activity**

There is also one final effective weapon in the defendant’s arsenal to suggest that a lower standard of care might necessarily be owed. It is worth consideration here as it has very recently been shown to be effective in cases of this type. In *Tomlinson v. Congleton Borough Council* Lord Hoffman famously stated that in order to assess a standard of care the factors that must be taken into account are: the likelihood of the injury occurring; its likely severity if it did occur; the cost of taking preventative measures; and the social value of any activity that would be prevented if steps to remedy the risk were taken.\(^{41}\) In that case, and also subsequently in *Trustees of the Portsmouth Youth Activities Committee (A Charity) v. Poppleton*,\(^ {42}\) it has been held as a useful additional factor for assessing the standard of care in occupier’s liability cases where responsibility for the welfare of the claimant has not been taken. Indeed in *Poppleton* it was suggested that had responsibility for Mr Poppleton’s safety been taken by an instructor that the standard of care may have been applied differently. In the interesting case of

\(^{41}\) [2003] EWHL 47, [2004] 1 AC 46, at paragraph [35].

Macintyre v. Ministry of Defence\(^{43}\) it was suggested that this argument could be applied where instructors had, as part of standard operating procedure, deliberately exposed clients to risk. In that case two teams of MOD instructing officers were leading recruits in the Alps on adventurous training. They were forced to cross a section of ground on the climb made up of very loose rock. The two leaders did so together, trailing the ropes behind them. A rock was dislodged which seriously injured the claimant. In coming to his conclusion, Spencer J applied Lord Hoffman’s test, including the social benefit of the activity. He cited the fact that this case was “After all ... adventurous training”\(^{44}\) and recognised the social value of a real but minimised risk in these kinds of activities.

Spencer J applied section 1 of the Compensation Act 2006 in coming to his decision, which, although it remains in most cases merely a codification of the common law, allows judges to take into account the beneficial nature of the activity that would be prevented in all cases when assessing an appropriate standard of care. Such an argument was also successful where the case of Uren v. MOD at first instance in the High Court. In giving his judgment, Field J stated, “a balance has to be struck between the level of risk involved and the benefits the activity confers on participants and thereby on society generally.”\(^{45}\) It is submitted that where the activity is inherently risky, this argument will prove to be particularly strong. The more likely it is that imposing a standard of care will have a detrimental effect, the more weight the argument is likely to carry.\(^{46}\) It is as yet unknown if such an argument would be effective in a case involving a civilian deliberately exposed to risk. It is submitted that given the negative attitude of the courts towards a risk adverse society that such an argument could potentially be effective. As Jackson LJ stated in his dissenting judgement in Barnes v. Scout Association, “It is not the function of the law to eliminate every

\(^{44}\) Macintyre v. Ministry of Defence [2011] EWHC 1690 (QB), [2011] All ER (D) 65 (Jul), Per Spencer J at paragraph [122].
\(^{45}\) Uren v. Corporate Leisure and another [2010] EWHC 46 (QB) per Field J at paragraph [59].
\(^{46}\) Dancing Around the Maypole: Doe s.1 of the Compensation Act 2006 Protect Beneficial Activities, Toby Gee.
iota of risk or to stamp out socially desirable activities.”

Although ultimately the court decided that case in favour of the claimant, a scout who had run in to a wall whilst playing an organised game in the dark, the sentiments expressed by Jackson were generally agreed with. A slight challenge would remain in presenting the activity as socially desirable, this should pose no great obstacle even for cases involving civilians. As the BMC states of climbing and mountaineering, these activities are “physically and mentally demanding as well as incredibly fulfilling.”

The same is true of all the sports focussed on in this article. Promoting physical and mental wellbeing and an ability to appreciate real risks seems likely to be considered a beneficial activity and one that the courts should support. This is despite the fact that such activities involve risks that are difficult to predict and to calculate.

Conclusion – a shifting standard

It is imperative that those acting for either claimants or defendants in cases involving supervised extreme sports realise that the cases, while rare, are complex. They require a careful approach taking account of a large number of factors, and must be considered in light of the in-built risks present in these activities. In assessing the standard of care one must consider first the basic factors: the severity of the potential injury, the likelihood of that injury, the costs required to prevent such an injury occurring and the social damage that may be caused by doing so. In addition to this, one must also consider industry practice in terms set out by the Bolam test, before considering whether the risk taken foreseeably, in all probability, materially increased the risk that the client was exposed to in any event. Experts should be consulted on these terms in order to ensure that appropriate advice is given. While much has been made of consenting to the risks posed by these activities such an argument should only be effective where those risks were obvious or otherwise understood by the client. Finally the nature of the clients themselves will affect the standard of care they were owed. Important factors seem from the case law to include: age,


military or civilian, experience and their purpose in doing the activity. In this way assessment can be made on a case-by-case basis.

The assessment and overlaps between these factors are affected significantly by public policy. In the mid-nineties, following increased regulation, courts seemed far less willing to accept that those participating in these activities could be exposed to risks by their instructors. Following Tomlinson and the line of cases that have come after it would appear that the weight given to the factors mentioned is shifting. While the law still attempts to steer a course between a paternalistic attitude and allowing personal autonomy, the way in which this balance is assessed seems to have changed. More focus is being put on a balance between the foreseeable and probable increase in risk and, on the other hand, the risk of preventing a socially beneficial activity. In turn, less focus seems to be put on the balance between a modified Bolam test (albeit with some recent exceptions), and the consent of the client. Given recent case law and the decline in regulation of these activities it is likely that further shifts may well occur in the near future.

FREDERICK LYON
Has Bolitho redressed the balance between defendants and claimants in cases of clinical negligence?

CELYN COOPER

Abstract
For the purposes of discerning whether the decision in Bolitho\(^1\) has redressed the balance between defendants and claimants in cases of clinical negligence, I will proceed to give an introduction to negligence as tort before progressing to the emergence of professional, clinical negligence. Subsequently, I will examine the case of Bolam\(^2\) which gave rise to the Bolam Test and identify its shortcomings which will allow me to analyse which, if any, the decision of Bolitho\(^3\) has rectified. Given the wide scope of the question, bestowed upon me with gratitude, and the broad ranging areas of clinical negligence where the Bolam Test has affected, for the purposes of the essay I will primarily concentrate on its effects on the Standard of Care needed to satisfy a Breach of Duty.

Perspective
“Brain-damaged boy awarded £6.4m”,\(^4\) “Mothers put at risk by inferior overnight care, maternity chief warns”\(^5\) and “The “catalogue of errors” that cost this father his life”.\(^6\) These are just a handful of illustrations of the increasingly litigious relationship between the patient and medical professional in the 21\(^{st}\) Century. Whether it emanates from a failure to provide adequate care, clinical errors or by simply failing to obtain sufficient consent, society has engendered what is proverbially referred to as Britain’s “no-win no-fee” culture of blame. Something, I very much doubt, Lord Atkin envisaged when he gave the landmark ruling of Donoghue v Stevenson\(^7\) which inaugurated the modern law of negligence almost a century ago.

---

1 Bolitho v City and Hackney Health Authority [1997] 4 ALL E.R. 771.
2 Bolam v Friern Hospital Management Committee [1957] 1 W.L.R. 582.
3 Bolitho v City and Hackney Health Authority [1997] 4 ALL E.R. 771.
7 Donoghue v Stevenson [1932] UKHL 100.
Introduction

The Tort of Negligence

“Discernable evidence of the emergence of negligence as a separate tort” in its own right emanated from 1825 onwards. Following a generation of the application of a writ of trespass against personal injury, the judiciary endeavoured to differentiate between deliberate and inadvertent harm. It proceeded as a reflection of a period where the industrial revolution generated an “increased risk of injury to those working in factories, mines, quarries, and other dangerous situations” and it was a reflection entrenched first by the enactment of the Workmen’s Compensation Scheme in 1897. The employer would owe a duty of care to his employee where he could establish that the harm was caused during his period of employment.

As a result, the precedent set by Lord Atkin in *Donoghue v Stevenson* in 1932 heralded a paramount shift in how and where a duty of care could arise. Where a duty of care owed could only give rise to a claim of negligence when an exceptional contractual relationship existed between the parties, it now followed that the duty will automatically arise in such circumstances where there is no established duty of care, where an individual fails to avoid “acts or omissions which are reasonably foreseeable” that would have a profound effect on those “so closely and directly affected” by such an act. This “neighbour principle”, as the statement is often referred to, singlehandedly established a firm, new category of duty between the manufacturers of goods and their potential consumers (thus providing a remedy to their suppliers), and finally allowed for citizens to bring claims of negligence in any contingency where the criteria for establishing the duty could be met. That is, the claimant must establish that the

---

11 Donoghue v Stevenson [1932] UKHL 100.
12 Lord Atkin’s Neighbour Principle.
defendant owed him a duty of care, of which was breached and inevitably led to cause the “legally recognised”\textsuperscript{13} harm he suffered.

The post-War political and social shifts of the 1940s saw the presiding Government assume responsibility over the welfare of the British citizens through various enactments of legislation which led to the creation of the Welfare State and the National Health Service. Combined with the ruling of \textit{Donoghue v Stevenson}\textsuperscript{14} it illustrated “accordingly, [that] carelessly inflicted harm [now] occupies a central position in the English law of torts, and it is [now]...recognised that, in certain circumstances, persons guilty of careless conduct should be liable to their victims”,\textsuperscript{15} But what of the evolution of clinical negligence? With the state funded NHS, “Tort law was no longer the primary support for those suffering loss”.\textsuperscript{16}

\textbf{Clinical and Professional Negligence}

The developments following the founding concept of the Welfare State and the establishment of the National Health Service in 1948 changed the face of “state medicine, welfare and public health”\textsuperscript{17} dramatically. It subsequently changed the way we as patients, see the medical profession as providers and arbiters of our healthcare. Doctors and health professionals enjoyed the privilege of standing so close to the state\textsuperscript{18} – incubating the “doctor-knows-best” paternalism and gained a favoured position of circumstance in the Courts where the law increasingly set doctor's negligence apart “from regular trade negligence”.\textsuperscript{19} Unwavering faith of their profession meant that they could potentially exercise their clinical discretion to the further detriment of the health of their patients without great

\textsuperscript{14} Donoghue v Stevenson [1932] UKHL 100.
\textsuperscript{15} Murphy, J. "Street on Torts", 12th Edition, Oxford University Press, 2007 at page 23.
\textsuperscript{16} Lord Atkin’s Neighbour Principle.
\textsuperscript{17} \url{http://www.thelancet.com/journals/lancet/article/PIIS0140-6736(10)60081-5/fulltext} accessed 31st January at 16:47pm.
\textsuperscript{18} \url{http://www.thelancet.com/journals/lancet/article/PIIS0140-6736(10)60081-5/fulltext} accessed 31st January at 16:47pm.
\textsuperscript{19} \url{http://www.thelancet.com/journals/lancet/article/PIIS0140-6736(10)60081-5/fulltext} accessed 31st January at 16:47pm.
judicial impedance. Illustrative of this, could be the figure\textsuperscript{20} of compensation claims as means of settling medical negligence which was “handed out” in 1952 as compared to 2000: £34,472 to nearly £78,000,000. However, the question of whether this figure is illustrative of the Courts unwillingness to support a litigant’s claim of negligence against a doctor, or merely the reluctance of patients harmed to bring a claim against him is an evocative one to take into account.

The greatest support, however, to upholding the infallibility of medical professionals came from \textit{Bolam v Friern Hospital Management Committee} [1957]\textsuperscript{21} where McNair J unequivocally empowered doctors to govern the standard of care and competence where a claim of negligence arises, which weighted heavily against any claimant.

\textbf{Bolam v Friern Hospital Management Committee [1957]\textsuperscript{22}}

Establishing the existence of a duty of care between doctor and patient is evidently unproblematic as there is sufficient proximity between the two parties, and it is a duty imposed which is universally “fair, just and reasonable” in the circumstances as purported by the ruling in \textit{Caparo Industries plc v Dickman} [1990]\textsuperscript{23} The duty of care between doctor and patient is in effect, implied. Subsequently, the first hurdle for the aggrieved patient to overcome is to establish the breach of duty. That is, the medical professional was negligent: a hurdle theoretically impossible to cross following the decision in \textit{Bolam}.\textsuperscript{24}

The facts of \textit{Bolam}\textsuperscript{25} are well known to all. Having been subjected to electro-convulsive therapy without the administration of a muscle relaxant drug or method of restraint, the claimant, Mr. John Bolam suffered from a fractured hip. When a claim of professional negligence was brought before the High Court,
expert medical opinion differed as to the practice of electro-convulsive therapy in such a manner.

A number of issues must be raised before we proceed to explore the Bolam Test itself. The case posed a number of difficulties for McNair J at the Queen's Bench Division, namely for the fact that there had been no prior precedent set on medical negligence. This is evident from his reference to the case of Hunter v Hanley [1955], a Scottish case to which he relied heavily on the principle set down by Lord President Lord Clyde: “The true test for establishing negligence...on part of the doctor is whether he has been proved to be guilty of such failure as no doctor of ordinary skill would be guilty of if acting with ordinary care”. It seems a relatively reasonable principle to follow, for the professional should always be judged against the professional of equal competence. It’s simply unfeasible to assume a duty of care of the same standard between the man on the Clapham Omnibus and the practicing Cardiologist for example. This became known as the first limb of the Bolam Test and is relatively uncontroversial: “The test is the standard of the ordinary skilled man exercising and professing to have that special skill”. The second case McNair J refers to in his judgement, however, could provide the rationale for the decision to which many claimed placed an undue restriction on claimants and has had a profound effect on the fraught area of medical negligence as a whole: the second limb of the Bolam test.

Roe v Ministry of Health [1954] 2 ALL ER 131 is a case which optimises the position of doctor’s in this period. Combined with Woolley v Ministry of Health [1954] 2 ALL ER 131 the Court of Appeal ruled that the anaesthetists “misadventures” which led to the paralysis of both claimants were not negligent. Lord Denning declared that "medical science has conferred great benefits in mankind, but these benefits are attended by considerable risks...
cannot take the benefits without taking risks...Doctors, like the rest of us, have to learn by experience; and experience often teaches in a hard way”. Denning confers a sense of humanity into the profession which inadvertently evokes a sense of compassion for what has occurred. It is, to some extent, a common sense judgement. But, to his detriment, he concludes that "we should be doing a disservice to the community at large if we were to impose liability on hospitals and doctors for everything that happens to go wrong”. I am assured to assume that he did not intend this comment literally, but at the mercy of McNair's judicial discretion, it may have been used as a means of protecting the medical profession without the “proper sense of proportion require[d]...to the insist[ence] on due care for the patient at every point”.

Consequently, following the respective judgements of Lord Denning and Lord Clyde, McNair J held that Friern Management Committee were not negligent in relation to the harm caused to Mr. Bolam. The second limb of the Bolam test centred on conflicting medical evidence, and it often “centres on just what does constitute “proper practice” or “ordinary competence”. It was thus famously held that “a doctor is not guilty of negligence if he has acted in accordance with a practice accepted as proper by a responsible body of medical men skilled in that particular act...a doctor is not if he is acting in accordance with such a practice, merely because there is a body of opinion taking a contrary view”. What this statement subsequently gave rise to, was a judicial perception “that all Bolam requires is that the defendant fields of experts from his or her medical specialty prepared to testify that they would have followed the same course of management...as the defendant”. That is, it was for the medical profession and not the judiciary to determine the standard of care expected of medical

31 Roe v Ministry of Health [1954] 2 ALL ER 131 per Lord Denning at 137.
32 Roe v Ministry of Health [1954] 2 ALL ER 131 per Lord Denning at 137 at 139.
33 Roe v Ministry of Health [1954] 2 ALL ER 131 per Lord Denning at 137 at 139.
35 Bolam v Friern Hospital Management Committee [1957] 1 W.L.R 582 at 122.
practitioners by law, and it is their duty to “measure...whether one has discharged his or her standard of care in the management of the patient”.

Expert opinion become conclusive evidence in doctors’ denial of medical malpractice, and following Bolam’s approval in both the House of Lords’ decisions in Whitehouse v Jordan and Maynard v West Midlands it became “the touchstone of liability for negligence”. By allowing doctors to set the standard of care in negligence, and the judiciary professing an “obvious reluctance to question the evidence of a “responsible” medical practitioner”, any possibility for the patient to successfully establish the breach of professional duty was severely diminished.

The Implications of Bolam

Per McNair J, “If the result of the evidence is that you are satisfied that his practice is better than the practice spoken of on the other side, then it is really the stronger case”. It is unlikely that our honourable friend intended to entrust to the medical profession “the standard of care expected of a doctor by law” which would be conclusive, and certainly having been aware at the time of the increasing disuse of juries in medical negligence cases evidently left open the potentiality “of the court having a more active role in setting the standard...by objectively evaluating the practice proffered by each of the parties”.

---

38 Bolam v Friern Hospital Management Committee [1957] 1 W.L.R 582.
40 Maynard v West Midlands [1984] 1 WLR 634.
43 Bolam v Friern Hospital Management Committee [1957] 1 W.L.R 582.
44 Bolam v Friern Hospital Management Committee [1957] 1 W.L.R 582.
Evidently, the mere mention of invoking the Bolam Test has become “enough to defeat claims sufficiently contestable to reach the courts”.47 The test indefinitely rests for the benefit of the defendant, and not the claimant. By taking a sociological, rather than normative approach to questions of negligence, that is, favouring “what is done in the course of standard practice by clinicians”,48 it “deprived the courts of the opportunity of precipitating change where required in professional standards”49 because it conceives that the medical profession “sets for themselves the legal standard by eliciting support of a responsible body of medical opinion”.50 From being “over protective and deferential towards doctors” to “a view that Bolam did not necessarily protect the community against unsafe medical practices”,51 then why, if this is the case, did the House of Lords approve a first instance decision, with no apparent authority and “certainly not binding”52 by virtue of Whitehouse53 and Maynard?54 Evidently, “Bolam, in suits for medical malpractice, was interpreted, whatever the judges might say, to allow judgement by colleagues to substitute for judgement by the courts”.55

The House of Lords approval has proven detrimental to any pursuit of medical negligence by a claimant. Lord Scarman unequivocally declared the Court’s objection to scrutinising “the evidence of a “responsible” medical practitioner”56 in Maynard:57 “a judge’s preference for one body of distinguished professional opinion to another is not sufficient to establish negligence in a practitioner whose actions have received the approval of those opinions, truthfully

54 Maynard v West Midlands RHA [1985] 1 ALL ER.
57 Maynard v West Midlands RHA [1985] 1 ALL ER.
expressed, honestly held, were not preferred”.58 While there will always be competing opinions by medical practitioners in the exercise of their clinical discretion, placing the threshold so unbelievably high at the preliminary stages of litigation, while thwarts any floodgate argument of future undignified claims of negligence, is usually at the expense of fairness, and justice for honest claimants. It’s simply not a cogent justification for employing a stance where the judiciary are in effect, “dictated to”59 and do not “appraise what would be reasonable under the circumstance”60 and certainly do not “state the expected standard [of care needed] to define the boundaries of reasonable conduct”.61

What is certainly disheartening, is the thought that the Bolam “defence” could have been completely departed from had the decision in Hucks v Cole62 been reported earlier.63 Certainly, Sachs LJ sitting in the Court of Appeal had provided an “example of judicial independence”64 for the House of Lords to consider in both instances of Maynard,65 Whitehouse66 as well as Sideaway,67 departing away from applying the Bolam Test in almost unwavering anticipation that McNair J’s declaration could move patient’s rights from “the centre of medical negligence”.68 Not only did he declare the inconclusivity of the “body of professional opinion sanctioning the defendant’s conduct”69 in relation to establishing the standard of care, he also subjected the expert opinions to rigorous scrutiny, which we shall see emanating through anachronistic echoes in Lord Browne-Wilkinson’s

---

65 Maynard v West Midlands RHA [1985] 1 ALL ER.
67 Sidaway v Board of Governors of the Bethlehem Hospital and the Maudsley Hospital [1985] 1 ALL ER 643, HL.
68 http://www.thelancet.com/journals/lancet/article/PIIS0140-6736(10)60081-5/fulltext
judgement in Bolitho.\(^{70}\) "...the fact that other practitioners would have done the same thing as the defendant practitioner is a very weighty matter to be put on the scales on his behalf; but it is not...conclusive. The court must be vigilant to see whether the reasons given for putting a patient at risk are valid...I was in the end fully satisfied that...[the failure to act] was not merely wrong but clearly unreasonable. The reasons given by the four experts do not to my mind stand up to analysis...the fact that others say they would have done the same neither ought to nor can in the present case excuse him in an action for negligence".\(^{71}\)

If it were an area, any other than clinical negligence, where a professional's conduct was called into question, the Court would have no hesitation in subjecting expert opinion to rigorous scrutiny, even where the alleged negligent act “conformed to the general practice of the profession”.\(^{72}\) This is evident from the Privy Council’s decision in Edward Wong Finance,\(^{73}\) where it was held that “a practice nearly universally enforced by solicitors in Hong Kong was nonetheless found to be negligent”\(^{74}\) because “evidence that other respected professionals followed the same unsafe practice was not sufficient to amount to conclusive evidence”.\(^{75}\) It is often argued that this decision is incomparable to those we find in cases of medical negligence, because it is a profession about which the judiciary have inherent, in-depth knowledge. While Bolam does relieve the judiciary from the burden of complex, intricate thoughts of medical practice, it should not sanction negligence through simple peer evidence\(^{76}\) to enjoy that relief which is at the expense of the aggrieved claimant. Of the six claims of medical negligence that have reached the House of Lords in 16 years, the claimant has failed in all.\(^{77}\) The Royal Commission on Civil Liability and

\(^{70}\) Bolitho v City and Hackney Health Authority [1997] 4 ALL E.R. 771.

\(^{71}\) Bolitho v City and Hackney Health Authority [1997] 4 ALL E.R. 771 per Lord Browne-Wilkinson at 243.


\(^{73}\) Edward Wong Finance Co. Ltd v Johnson Stokes & Master (a firm) [1984] A.C. 296.


Compensation for Personal Injury\textsuperscript{78} in 1978 effectively “expressed concern that very few claimants in medical negligence cases, as compared with all other types of negligence cases, were successful”.\textsuperscript{79}

“The Bolam Defence had the effect of 40 years of providing strong protection for doctors accused of negligence”,\textsuperscript{80} and certainly created a situation in which claimants appeared to have the balance in clinical negligence cases weighted against them. We have seen its birth at First Instance, its approval at the House of Lords at the expense of ignorance to both Court of Appeal and Privy Council authorities; while not binding, persuasive enough to attach greater weight to them than McNair’s judgement. It seems that the judiciary have had complete disregard of their own professional discretion in adopting approaches supported and found in eminent authorities; approaches which could have easily redressed the imbalance facing claimants. 40 years respectively, came the “eureka!"\textsuperscript{81} moment we had all been waiting for, in light of the ruling of Bolitcho – but has it been enough to articulate the doctors duty of care in terms of patients rights?\textsuperscript{82}

\textit{Bolitho v City and Hackney Health Authority [1997] 4 ALL E.R. 771}

Towards the end of the 20\textsuperscript{th} century, and certainly by the beginning of the 21\textsuperscript{st} century, there were “greater calls for patient’s rights”\textsuperscript{83} and a distinct “shift in the rhetoric of medical negligence cases favouring a more interventionist stance”\textsuperscript{84} emerged. A number of appellate decisions\textsuperscript{85} provided “clear signs that judicial perceptions of reasonableness...are relevant to the determination of the requisite standard”.\textsuperscript{86}

\begin{itemize}
  \item \textsuperscript{78} The Pearson Commission.
  \item \textsuperscript{79} Harwood, V. “Clinical Negligence”, Ex Abundante 2010 at page 70.
  \item \textsuperscript{80} Harwood, V. “Clinical Negligence”, Ex Abundante 2010 at page 73.
  \item \textsuperscript{81} Grubb, A. "Negligence: Causation and Bolam", 1998 6 Med. L. Rev. 378 at page 380.
  \item \textsuperscript{83} http://www.thelancet.com/journals/lancet/article/PIIS0140-6736(10)60081-5/fulltext
  \item \textsuperscript{84} Teff, H. “The Standard of Care in Medical Negligence – Moving Away from Bolam?” Oxford Journal of Legal Studies, 1998 , 18(3): 473-484 at page 477
  \item \textsuperscript{85} Loveday v Renton [1990] 1 Med LR 117; Knight v Home Office [1990] 3 ALL ER 237; Joyce v Merton, Sutton and Wandsworth Health Authority [1996] 7 Med LR 1
\end{itemize}
The facts of Bolitho, and the death of 2 year old Patrick are one of the most harrowing in the literature. Having been admitted to hospital suffering from respiratory complications, his “condition deteriorated and he suffered a cardiac arrest, leading to brain damage and subsequently his death.” While the case hinged on causation, and whether the paediatric registrar exercised her duty negligently in failing to “arrange for [a] prophylactic intubation”, a crucial “judicial statement was made upon when the Bolam principle could be challenged”. 

Having finally paid due regard to the eminent ratio’s in Edward Wong and Hucks, the House of Lords ruled that the court could assess and scrutinize whether a body of professional opinion “[was] reasonable or responsible” and that it would be the arbiter of the standard of care in negligence, “despite expert evidence agreeing with the defendant’s course of action”.

"...the court is not bound to hold that a...doctor escapes liability for negligent treatment...just because he leads evidence from a number of medical experts who are genuinely of opinion that the defendant’s treatment...accorded with sound medical practice." In overt and express reference to Wong and Hucks, Lord Browne-Wilkinson declared that they demonstrated that “despite a body of professional opinion sanctioning the defendant’s conduct, the defendant can properly be held liable for negligence...because, in some cases, it cannot be demonstrated to the judge’s satisfaction that the body of opinion relied upon is reasonable or

---

87 Bolitho v City and Hackney Health Authority [1997] 4 ALL E.R. 771.
91 Edward Wong Finance Co. Ltd v Johnson Stokes & Master (a firm) [1984] A.C 296
responsible". Furthermore, at 243 Lord Browne-Wilkinson echoes Sachs LJ’s dicta in Hucks by introducing the concept of balancing risks and benefits, that is, “the judge must assure himself that experts have directed their minds to the question of comparative risks and benefits and have reached a defendable conclusion on the matter”.101

The Implications of Bolitho – Does it Redress the Balance?
This new robust approach to questioning professional medical opinion provides and bestows at least some protection for a claimant; in “reducing the risk of legitimating the lowest common denominator of accepted practice” by medical practitioners. By applying the Bolam Test, Lord Browne-Wilkinson was able to address the decades of its misinterpretation, and invites future Justices, when faced with similar cases (though such circumstances will be rare) to “subject the opinion [in question] to logical analysis and critical evaluation [until they are] satisfied that it withstands such tests”. That is, “the views of expert witnesses must not only be honestly...held, but must also be defensible”. As a re-iteration, McNair J never intended for medical opinion to be conclusive: “...that does not mean a medical man can obstinately and pig-headedly carry on with some old technique if it has been proved contrary to what is really substantially the whole of informed medical opinion”.106

By effectively applying this level of “scrutiny”, with the burden still resting on the claimant to establish that the defendant breached his duty, it means that a burden, in return, will have to be placed on the expert for his opinion to withstand judicial scrutiny. The weight could be said to be equally balanced

98 Bolitho v City and Hackney Health Authority [1997] 4 ALL E.R. 771.
99 Bolitho v City and Hackney Health Authority [1997] 4 ALL E.R. 771.
102 Bolitho v City and Hackney Health Authority [1997] 4 ALL E.R. 771.
106 Bolam v Friern Hospital Management Committee [1957] 1 W.L.R 582 at 587.
between the parties. It serves as a “reminder to the courts of the need to retain control over the legal issue of the standard of care required of professionals”. While Lord Browne-Wilkinson respectively held the foresight to state that such cases concerning the breach of duty of medical professionals will be rare, there have been a small number of authorities decided which illustrate that the judiciary have avidly “challenged the logic or defensibility of medical evidence”. Reynolds, Marriott, Antoniades and Richards all illustrate that the judiciary are now prepared and “use the same mixture of common sense and logical analysis [used] to scrutinise other expert evidence in negligence claims against profession”.

Bolitho exemplifies the rise in the acceptance of patients rights in the medical sphere, and as the Honourable Lord Woolf himself stated, in and amongst the birth and growth of excessive litigation “the courts became increasingly conscious of the difficulties which bona fide claimants had in successfully establishing claims” as only 17% of them were successful. Combined with the incessant damnations of the Bolam Test, it was only a matter of time before the halo was removed from the automatic presumption of beneficence of medical practitioners thus giving claimants a “fighting chance” to hold them to account to the grievance they’ve caused. But is it a mere narrow exception to the general applicability of Bolam?

114 Bolitho v City and Hackney Health Authority [1997] 4 ALL E.R. 771.
The ruling in *Bolitho*\(^1\) was simply a “modification of the *Bolam* test rather than a radical departure from it”.\(^2\) Because they are now, effectively applied simultaneously, it is difficult to assess whether *Bolitho*\(^3\) has redressed the weight balanced against claimants in clinical negligence cases. While *Bolitho*\(^4\) grants more “powers”, or authorizes greater judicial scrutiny of the evidence of medical experts, the paternalism is still upheld in the form of the view that to question eminent medical opinion in itself is illogical. Take the case of *Wisniewski*\(^5\) as an example. The Court of Appeal rejected the trial judge’s preference for the claimant’s expert testimony on the basis, just as Lord Browne-Wilkinson had advocated in *Bolitho*, of a “comparative risk/benefit analysis”\(^6\) and concluded that it was “quite impossible”\(^7\) to find the defendant’s “eminent consultant[s] views as [i]logical and [u]nsupported by responsible doctors”\(^8\). On closer inspection of the selective words used by Lord Browne-Wilkinson in *Bolitho*,\(^9\) it does bestow some credibility to the view that *Bolitho* is merely a “gloss”\(^10\) of the *Bolam* test: “it would very seldom be right for judge to reach the conclusion that views genuinely held by a competent medical expert are unreasonable”.\(^11\) The Court of Appeal’s seems almost disgusted at the thought of seeing another professional being discredited for his own views and it certainly, and surprisingly, echoes greater the judicial attitudes seen pre-*Bolitho*.\(^12\) In particular, Lord Scarman’s “castigation of a trial Judge for presuming to prefer one body of distinguished professional opinion to another”.\(^13\)

---

Furthermore, the fact that the Bolitho framework is only engaged as a “trump”\textsuperscript{132} where the defence produces evidence (for the purposes of satisfying the Bolam threshold) which is “unreasonable in the Bolitho sense” it provides little impressive hope for the claimant in succeeding.\textsuperscript{133} As Mulheron eloquently expresses, “...once Bolam applies, the mere fact of differences in expert opinion cannot lead to a rejection of Bolam evidence, as Bolam itself acknowledged (a man is not negligent if he is acting in accordance with such a practice, merely because there is a body of opinion who would take a contrary view”).\textsuperscript{134}

The staunch supporters of the Bolam\textsuperscript{135} test, in light of the decision in Bolitho argued that Bolam did (properly applied) “balance the rights” of competing claims: “it protects doctors who act in accordance with the provisions accepted by their professions and it allows a patient to sue”.\textsuperscript{136} Bolitho is thus coercive; prescribing the exercise of clinical discretion to defensive practice which would have a profound effect on the doctor/patient relationship as well as encouraging excessive litigation.\textsuperscript{137} To them, Bolitho\textsuperscript{138} is too objective, allowing the judiciary to construe their own opinion without adequately relying on the medical testimony. Even recent case law suggests that there will rarely be “grounds for invoking Bolitho”,\textsuperscript{139} and where it was invoked in that instance, the Court of Appeal overruled the trial judge’s decision to find the GP negligence as he had “impose[d] his own opinion, regardless of the practice of the medical profession”.\textsuperscript{140} I inherently disagree with the majority of commentaries which suggests that the Courts are ill equipped to answer questions couched in medical

\begin{flushright}
\textsuperscript{135} Bolam v Friern Hospital Management Committee [1957] 1 W.L.R 582.
\textsuperscript{138} Bolitho v City and Hackney Health Authority [1997] 4 ALL E.R. 771.
\textsuperscript{140} Ministry of Justice v Carter [2010] EWCA Civ 694 per Sir Scott Baker at 28 and 29.
\end{flushright}
principle, terms and practices. The developed process and growth of judicial review, and the judiciary’s oath to remain impartial and unbiased in all instances brought before them show that illustrate that they are capable of adopting “a more proactive approach to resolving conflicts”\(^{141}\) between parties.

**Conclusion**

The decision in *Bolitho*\(^{142}\) should be welcomed, in spite of *Bolam’s*\(^{143}\) 40 year reign and should, to all intents and purposes, serve to redress the imbalance claimants faced when bringing actions of clinical negligence against a medical practitioner. Certainly, there are cogent arguments to suggest that it may not fulfil its unmitigated potential. The judiciary, it seems, create self-inflicted hierarchal obstacles for themselves, firstly by disregarding eminent precedent (as in *Edward Wong*\(^{144}\) and *Hucks*)\(^{145}\) and secondly, for refusing to be seen overruling any preceding precedent or views set.

Then what options are left? The use of clinical guidelines, it has been said, could “enable the court to utilise the *Bolitho* principle more proactively in setting the expected standard of care required of doctors in cases of medical litigation”.\(^{146}\) By providing the judiciary with this aid of interpretation, “material independent of the particular dispute before him”,\(^{147}\) it will allow them to assess the expert opinion “in context of the profession’s considered judgement”.\(^{148}\) While guidelines aren’t strictly mandatory, they serve a persuasive purpose and could aid a claimant wishing to establish a breach of duty. In *Richards (a child) v Swansea NHS Trust*\(^{149}\) Field J expressly referred to NICE Guidelines on Caesarean

---


\(^{142}\) *Bolitho v City and Hackney Health Authority* [1997] 4 ALL E.R. 771.

\(^{143}\) *Bolam v Friern Hospital Management Committee* [1957] 1 W.L.R. 582.

\(^{144}\) *Edward Wong Finance Co. Ltd v Johnson Stokes & Master (a firm)* [1984] A.C 296


\(^{149}\) *Richards (a child) v Swansea NHS Trust* [2007] EWHC 487 (QB).
births in concluding that the medical practitioner owed the claimant a duty of care and was in breach. Furthermore, as noted by Brazier and Miola,\(^{150}\) the use and reference of guidelines in the courts, “coupled with the decision in Bolitho”\(^ {151}\) allows the judiciary to not only readily scrutinize expert opinion, but make them "better equipped to so".\(^ {152}\)

However, there is little use for the deployment of clinical guidelines in medical negligence cases where the guidelines themselves are effectively “created” by doctors. If the judiciary were to attach such great weight to them, the purpose of Bolitho would be undermined: doctor’s evidence would still be conclusive of establishing the standard of care. Furthermore, by “centrally direct[ing] standards, [which] will undermine the art of medicine”,\(^ {153}\) it stifles individual patient care: while clinical guidelines “are designed to promote best practice”,\(^ {154}\) it might not be best practice for that specific patient.

Greater hope and faith is given from the implementation of Lord Woolf’s review of the Civil Justice System, by virtue of the Civil Procedure Rules 1998. By virtue of Part 35 now states that where expert evidence is required, it will now be given, “whenever appropriate, by a single or joint expert. In addition all experts now owe their primary duty of the judge and not the parties”.\(^ {155}\) This has arguably “achieve[d] easier access to, and higher quality of, expert evidence”\(^ {156}\) which inhibits the majority of expert witnesses from modifying their opinions for


the purposes of satisfying the advocates that solicit their services, rather than the courts.\textsuperscript{157}

Evidently there are a number of factors, stereotypes and circumstances which contribute to the imbalance between a medical professional and patient in pursuit of a successful claim in negligence. \textit{Bolam}\textsuperscript{158} was, and still is, inherently inhibiting; curbing the courts from “exercising a restraining influence”.\textsuperscript{159} As Lord Woolf himself proclaimed, “it is unwise to place any professions or other body providing services to the public on a pedestal where their actions cannot be subject to close scrutiny”.\textsuperscript{160} While \textit{Bolitho}\textsuperscript{161} did redress the imbalance, to some extent, in permitting the Court to choose between the competing medical opinions, “in many instances, professional opinion will remain highly, and rightly influential”.\textsuperscript{162} We should not “sit back complacently”.\textsuperscript{163} With NHS compensation costs to rise to £807,000,000 within the next three years,\textsuperscript{164} and fewer than 4% are expected to come to court,\textsuperscript{165} it is difficult to foresee whether the utilisation of \textit{Bolitho},\textsuperscript{166} clinical guidelines and the arrogation\textsuperscript{167} of the judiciary into “making clinical judgements”\textsuperscript{168} will have any impact at all on a claimant’s successful claim of negligence. Certainly not while \textit{Bolam}\textsuperscript{169} remains to reign supreme.

\textsc{Celyn Cooper}

\textsuperscript{158} Bolam v Friern Hospital Management Committee [1957] 1 W.L.R 582.
\textsuperscript{161} Bolitho v City and Hackney Health Authority [1997] 4 ALL E.R. 771.
\textsuperscript{164} http://www.guardian.co.uk/society/2009/aug/19/nhs-legal-costs-compensation
\textsuperscript{165} http://www.guardian.co.uk/society/2009/aug/19/nhs-legal-costs-compensation
\textsuperscript{166} Bolitho v City and Hackney Health Authority [1997] 4 ALL E.R. 771.
\textsuperscript{169} Bolam v Friern Hospital Management Committee [1957] 1 W.L.R 582.
In the wake of the scandal surrounding the retention of organs at Bristol and Alder Hey hospitals in 1999, a spotlight was shone upon the Human Tissue Act 1961 (the 1961 Act) and its inadequacies were revealed. The chair of the Retained Organs Commission, Professor Margaret Brazier, went so far as to describe the 1961 Act as ‘a toothless tiger imposing fuzzy rules with no provision for sanctions or redress’, (M Brazier, ‘Organ retention and return: problems of consent’ 29 J.Med Ethics 2003; 29 30–33 at 31).

The exposure of practices at Bristol and Alder Hey made it clear to Parliament that the law had not kept pace with medical practice and society’s ethical expectations of that practice. Bristol and Alder Hey were found to be the tip of the iceberg. As Mr Paul Burstow, MP for Sutton and Cheam, informed the House of Commons on the 15th January 2004, ‘The chief medical officer's report on the census of organs and tissues retained by pathology services revealed that...poor practice and paternalistic attitudes were widespread. Over almost 30 years, 54,300 organs, body parts, stillbirths or foetuses were retained by pathology services...the census found poor cataloguing and recording of retained tissues and organs in hospitals and university research collections’, (Human Tissue Bill, HC Deb 15 January 2004 vol 416 c1006).

The 1961 Act had legitimised activities which not only shocked the general public but it was contested were in breach of the soon to be introduced Human Rights Act 1998 (HRA). Section 6(1) of the HRA enacted that ‘It is unlawful for a public authority to act in a way which is incompatible with a Convention right’ and yet the 1961 Act authorised the National Health Service to take and retain human tissue and organs from patients without consent being an absolute requirement; arguably in breach of an individual’s Convention for the Protection of Human Rights and Fundamental Freedoms (ECHR) Article 8 and Article 9 rights. ‘It is well established in ECtHR jurisprudence that the concept of ‘private
life’ includes a person’s physical and moral integrity; *X and Y v Netherlands* [1985] ECHR 4, [22]’, (Law and the Human Body: Property Rights, Ownership and Control; Rohan Hardcastle, 2009, p59).

Following the Redfern and Kennedy inquiries, the Chief Medical Officer, Sir Liam Donaldson, made 17 recommendations which the government accepted and in 2002 a consultation report ‘Human Bodies, Human Choices the law on human organs and tissue in England and Wales’ was published. The Human Tissue Bill was introduced into the House of Commons in December 2003, was brought to the House of Lords in June 2004 and received royal assent on the 15th November 2004. The Human Tissue Act 2004 (the Act) repealed not only the 1961 Act but also the Anatomy Act 1984, the Human Organ Transplants Act 1989 and the Human Tissue Act (Northern Ireland) 1962. The Act finally came into force in September 2006 after secondary legislation was written and the ‘Human Tissue Authority’ (the HTA) was established.

The first version of the Bill focussed on the requirement of informed consent, which was described as its ‘golden thread’. This ‘drew a vociferous response from the pharmaceutical, pathology, medical research and research charity communities in particular, alleging that vital research would be compromised’ (David Price, ‘The Human Tissue Act 2004’, The Modern Law Review (2005) 68:798-821).

During the 15th January 2004 second reading of the Bill, Dr Brian Iddon, MP for Bolton South East, described researchers as being ‘concerned that they may be about to commit a criminal act by examining tissue for research, when consent was not gained for such a research application’, (Human Tissue Bill, HC Deb 15 January 2004 vol 416 c992), and Dr Doug Naysmith, MP for Bristol North West, gave the opinion that, ‘Although it is clear that the public need to be confident that strict guidelines and procedures are in place to control the use of organs and tissues both from cadaveric and living donors—whether to be used for transplantation, teaching or research purposes—it is crucial that regulations and guidelines are not drawn so restrictively that they inhibit necessary and
worthwhile clinical medicine and research’, (Human Tissue Bill, HC Deb 15 January 2004 vol 416 c1034).

Amendments were made in response to these concerns, described by Lord Warner in the House of Lords debate on the 22nd July 2004 as, ‘important ones that recognised the practical aspects of the use of samples recovered during diagnostic and surgical procedures. I hope that medical research interests will recognise that we have responded to their concerns in a practical way while preserving the integrity of the new consent arrangements that the public wish to see in this legislation’, (Human Tissue Bill, HL Deb 22 July 2004 vol 664 c369).

But did these amendments preserve the integrity of the consent requirement? Or did they instead permit practice which still falls short of ensuring respect for the individual?

According to Dr Stephen Ladyman, MP for Thanet South, there was no pre-legislative scrutiny in the publication of a draft bill due to a lack of drafting resources (Human Tissue Bill, HC Deb 15 January 2004 vol 416 c1043) and the Bill was the subject of a great deal of lobbying; ‘46 groups of amendments were tabled in Standing Committee and 99 new clauses and amendments were proposed at Report stage’ (David Price, ‘The Human Tissue Act 2004’).

Was it the many subsequent amendments which led to the legislation being described as ‘both unhelpful and unwieldy’? (Samantha Taylor and David Wilson, ‘The Human Tissue Act (2004), anatomical examination and the importance of body donation in Northern Ireland’, Ulster Med J 2007; 76 (3) 124-126)

**The Meaning of ‘Appropriate Consent’**
The Act is made up of three Parts and seven Schedules which cover the regulation of cadaver material and tissue from living persons for therapeutic and non-therapeutic purposes. The breadth of its application to ‘relevant material’ is demonstrated in Section 53 where ‘relevant material’ is defined as ‘material, other than gametes, which consists of or includes human cells’. The
definition excludes gametes as the Human Fertilisation and Embryology Act 1990 governs activities relating to reproductive cells.

Section 1 constitutes a list of activities which are lawful if carried out with ‘appropriate consent’. Section 5 makes it a criminal offence to carry out a Section 1 activity without appropriate consent. Sections 2 and 3 purport to interpret the meaning of ‘appropriate consent’ but do not go so far as to define ‘consent’ itself. Ultimately the level of ‘consent’ required was left for the HTA to decide upon, the Act falling short of specifying the meaning of ‘consent’ other than to define it by reference to the person who may give it, i.e. the person concerned, their nominated representative or a person in a qualifying relationship to them immediately before they died. Further, the Act does not broadly stipulate the form of consent necessary to be ‘appropriate’, other than to require in Sections 2(4) and 3(3) that consent be in writing when given for public display or anatomical examination.

The HTA in its ‘Code of practice 1: Consent requirements’ (the Code) at paragraph 35 states that; ‘Consent may differ in its scope as it may be generic or specific’, at paragraph 56 that; ‘the information required and the manner in which consent is obtained and recorded may vary depending on the particular circumstances’, at paragraph 60 that; ‘consent is valid only if proper communication has taken place’, and at paragraph 63 that; ‘establishments should provide appropriate information on the activities for which they are seeking consent. The information might be in the form of leaflets or information sheets, or might be contained within the consent form’ (http://www.hta.gov.uk/legislationpoliciesandcodesofpractice). The HTA rightly ensures that there can be no presumed consent as communication is necessary, but there is a ‘woolliness’ to the Code in its acknowledgement that procedures may vary and that information may be provided in a variety of ways. At paragraph 62 a web link to the ‘National Research Ethics Service’ model consent forms and information sheets is provided, but there is no absolute requirement to use them.
In line with European Directive 2004/23/EC, the HTA's ‘Directions given under the Human Tissue Act 2004: 001/2006’ (the Directions) Section 14(b) requires establishments to provide information prior to the donation of human tissues and / or cells which; ‘cover at least the purpose and nature of the donation, its consequences and risks’ and at Section 15 it must be ensured that those giving consent ‘understood the information provided, had an opportunity to ask questions and been provided with satisfactory responses’. The Directions and the Code may both be interpreted quite differently, however, by establishments in fulfilling their obligations. As such, where one person may receive a written information pack and form to complete and be given the opportunity to ask a number of questions prior to making a decision, another may be asked for consent over the phone on the basis of what is quite minimal information. 'How much information should be given to the individual asked to decide whether they consent in order for their consent to be legally valid? ...it is widely argued that a person's consent is ethically meaningless unless the person has been given detailed information about the proposed conduct.’ (Beyond Bristol And Alder Hey: The Future Regulation Of Human Tissue, Kathleen Liddell and Alison Hall, Medical Law Review 2005.13(170).

Should the Act itself have gone further in defining ‘appropriate consent’ rather than leaving it to the HTA? Are the Directions and the Code enough to ensure that an individual is given the opportunity to make a fully informed decision? The lack of specificity in the Act or in the HTA’s provisions as to the process of obtaining legally ‘appropriate consent’ indicates that consent is obtained more for the establishments involved, to protect them from liability, than out of a genuine respect for the individual.

Exceptions to the Consent Rule
Section 11 of the Act provides that the consent rule does not apply to the activities of a Coroner and Section 39 exempts the police when keeping material in connection with a criminal investigation. Section 9 allows for the use of ‘existing holdings’ (relevant material held before the Act came into force) for any specified purpose other than anatomical examination without consent; a
practical decision given that ‘the 54,000 organs, body parts, stillborn and foetuses that were retained between 1970 and 2001 (were) just a small proportion of the total quantity of stored tissue specimens in the UK.’ (Beyond Bristol And Alder Hey: The Future Regulation Of Human Tissue).

Section 7 allows the use of material without consent where ‘it is not reasonably possible to trace the person from whose body the material has come’ or ‘where reasonable efforts have been made’ to gain consent without response and it is ‘for the purpose of obtaining scientific or medical information about him’ which may be relevant to another person.

More surprisingly, as Schedule 1 Part 2 only applies to deceased persons, consent is not required for the extraction, storage or use of relevant material taken from a living person for the purpose of clinical audit, education or training related to human health, performance assessment, public health monitoring or quality assurance. The Act’s ‘Explanatory Notes’ states; ‘these purposes are ones considered intrinsic to the proper conduct of a patient’s treatment...or necessary for the public health of the nation (public health monitoring and health-related education and training).’ It seems that consent by the patient to the procedure is deemed to cover these possible consequences.

As a direct result of the research community’s lobbying during the Bill’s passage through Parliament, Section 1(10) of the Act permits the storage and use of tissue from the living without consent where the researcher is not in possession of information that identifies the person from whom it has come and the material is used in a research project with ethical approval.

According to paragraph 43 of the Code; ‘ethical approval may only be given by a recognised research ethics committee’ which is either ‘established under and operating to the standards set out in the governance arrangements issued by the UK Health Departments’ or ‘an ethics committee recognised by the United Kingdom Ethics Committee Authority (UKECA)’. With 213 recognised research ethics committees in the UK today (http://www.npsa.nhs.uk), it is no wonder
that Professor Jean McHale of the University of Leicester, prior to the Act coming into force, questioned the effectiveness of such an approval system; ‘it may be questioned whether this will provide a sufficiently consistent approach...across the country – moreover it will mean that such approval decisions will have relatively low visibility. Is there a danger that such decisions will be subject to what is ultimately ‘routine’ approval?’ (The Human Tissue Act 2004: Innovative Legislation – Fundamentally Flawed Or Missed Opportunity?, Jean McHale, Liverpool Law Review (2005) 26: page 177).

Professor McHale also notes that there ‘may still be privacy interests engaged in this context if ‘privacy’ is interpreted as an individual’s interest in autonomy ...A person may be unhappy as to the use of their material in a particular research project because they are opposed to that type of research in general or to a research project on religious or cultural grounds. While some individuals may be perfectly happy for their ‘spare’ material to be used, it is surely a dangerous generalisation to assume that all will be’, (The Human Tissue Act 2004: Innovative Legislation – Fundamentally Flawed Or Missed Opportunity? page 177).

The research community rallied for this amendment to the Act, but at what cost? It is true that, ‘we all benefit from the existence of the social practice of medical research. Many of us would not be here if infant mortality had not been brought under control, or antibiotics had not been invented. Most of us will continue to benefit from these and other medical advances’, (My Body, Your Body, Our Bodies, Jonathan Herring, Medical Law Review 2007 15 (34)). But what about those individuals whose opinions and beliefs do not tally with this essentially enforced altruism? ‘A devout Sikh might retain an interest in even the smallest scrap of bodily material. Orthodox Jews seek means to store amputated limbs to be buried with them when they die...A host of reasons may describe why a particular individual may ascribe value to bodily material from the living’, (Respecting the Living Means Respecting the Dead too, Sheelagh McGuinness, Oxford Journal of Legal Studies 2008 28 (297)).
A person who has an operation and their cells are removed may not wish to take that material home with them, but would they necessarily consent to it being used in any of the 6120 research projects approved in England in 2010? (figure provided by Gill Habicht, Business Delivery Manager, National Research Ethics Service). 'the Government was undoubtedly very exercised by the prospect of researchers being tempted abroad if human tissue use was over-regulated here, jeopardising Britain’s status as a major research player. The interests of living patients as a result became slightly subjugated in the debate, with matters of principle being partially (and needlessly) compromised for utilitarian ends’ (David Price, 'The Human Tissue Act 2004').

The decision to remove the need for consent where material is anonymised may be seen to follow the case of R. v Department of Health Ex p. Source Informatics Ltd (No.1) Court of Appeal (Civil Division) [2001] Q.B. 424, where it was held that when patient information regarding prescriptions was anonymised it could be sold to pharmaceutical companies for marketing purposes. But in that case it was simply information - statistics on a computer being sold, here it is human tissue – relevant material - actual pieces of people. 'It would be a fairly simple solution to routinely ask the patient for consent for the use of surplus tissue at the time that they agree to a given surgical or diagnostic procedure.' (Speaking for the Dead: the Human Body in Biology and Medicine, D Gareth Jones and Maja I Whitaker, Ashgate Publishing 2009, page 56), but instead the Government bowed to the pressures of the researchers and gave carte blanche approval for the use of this tissue without consent. The decision tarnishes the golden thread of the Act and like its predecessor the 1961 Act, allows practice which arguably falls short of ensuring respect for the individual and breaches an individual's ECHR Article 8 and Article 9 rights.

**DNA Analysis and Organ Donation**

Section 45 of the Act creates an offence where a person 'has any bodily material intending that any human DNA in the material be analysed without qualifying consent'. Given the concessions made to researchers, it is a surprisingly strict rule. ‘The unexpected ramification of section 45 is that clinical geneticists will no
longer have the discretion to decide whether or not to analyse material containing DNA donated by a person …who refuses to allow it to be analysed for the benefit of another family member’, (Beyond Bristol And Alder Hey: The Future Regulation Of Human Tissue). Prior to the Act the common law’s approach to a conflict allowed for a balancing of interests, assessing competing rights and minimising privacy interference but ultimately allowing geneticists to analyse DNA in certain circumstances, without consent. ‘The fact that a particular person has tested positive for a mutation may be personal information; but the fact that a particular mutation is in the family is an essentially different kind of information which should not be regulated solely with regard for a proband’s preferences’ (Beyond Bristol And Alder Hey: The Future Regulation Of Human Tissue). The strictness of this rule may show a respect for the wishes of the individual from whom the material originated, but what about the relative who will not find out they have a genetic propensity to develop a form of cancer? In its inflexibility, does this rule not deny that relative respect for the value of their life? ‘The ethical and legal regulation of confidential information has been based on an assumption that the information is of significance only for the relevant patient. However in the case of genetic conditions, this is no longer true’, (My Body, Your Body, Our Bodies). Such information could prove critical in providing early treatment, yet Section 45 of the Act fails to acknowledge that such situations will call for a balancing of rights and instead provides a rigid statutory support for the individual from whom the DNA originated alone.

In regards to organ donation, the Act maintains an ‘opt-in’ system, despite discussion during the passage of the Bill and some favourable comments made in support of an ‘opt-out’ system. Kenneth Clarke, MP for Rushcliffe, told the House of Commons on the 15th January 2004 that; ‘the Spanish experience shows that a system of presumed consent can be operated in an entirely sensitive way, and that families can be consulted and their wishes considered, but on a different basis from that which obtains here. The Spanish system assumes at the outset that there are no strong objections, which has resulted in a rate of cadaver donations more than twice as high as ours. The change was not drastic, but it has proved to be very significant’, (Human Tissue Bill, HC Deb 15 January 2004 vol
416 c988). It seems that the catalyst for the Act’s creation, the retention of organs at hospitals without consent, ensured that an ‘opt-out’ system would never be palatable to Parliament.

Prima facie, an opt-in system is entirely respectful of the individual; if you do not wish for any part of you to be used in any way after you die, it is simply a case of not joining the ‘Organ Donor Register’ (www.uktransplant.org.uk) and ensuring that family are aware of your wishes, so that after your death they know not to give consent if approached as a person in a ‘qualifying relationship’. Where the system may be challenged as not fully respecting the individual is where a person wishes to donate an organ to another specific person after death.

‘Although we are allowed to decide for ourselves whether or not we want to be organ donors upon our death, in the event that we do, we cannot attach any condition to our ‘gift’ to society or specify a particular recipient, even if this is a person with whom we have a personal relationship. Instead, somehow or another, our donation slips straight into the net of public resource and impartial allocation...If, on the other hand, we are alive when we donate, we may then legitimately direct our donation (our gift) to a specified person with whom we do hold a relationship of some kind or another’, (Directed and Conditional Deceased Donor Organ Donations: Laws and Misconceptions, Antonia J. Cronin, Medical Law Review 2010 18 (275)). Thus, the Act differentiates between an individual’s right to choice in life and their choice after death.

During the House of Commons Standing Committee G debate on the 27th January 2004, Dr Stephen Ladyman, MP for Thanet South, said that; ‘the fundamental principle that we must apply to interpreting the Bill is that material provided by people from their own body is theirs to control, and they must consent to how it is used’, (HC Standing Committee G col 59 27th January 2004). Why then is an individual who declares in life that they wish to donate an organ to a particular person after their death unable to do so? It is a two tier system which lacks respect for the wishes of some, in life, after their death. ‘The Convention is not applicable to deceased persons but rights under article 8 might be invoked prior to death in respect of the subsequent treatment of the cadaver’ (David Price, ‘The
Human Tissue Act 2004’). The majority of organ donors will be content that their bodies are used to help any other person after their death, but there are those who will have a specific recipient in mind, (as per Laura Ashworth in 2008; http://news.bbc.co.uk/1/hi/england/bradford/7344205.stm), and unfortunately unless they begin the process of becoming a donor in life, they will not be permitted to make a direct donation in death.

**Criminal Penalties and Civil Remedies**

A key change introduced by the Act is the criminalisation of certain activities under Sections 5, 8, 25, 30, 31, 32 and 45, which may result in a term of imprisonment of up to 3 years, a fine, or both. Lord Warner, during the House of Lords debate on the 25th October 2004, declared that; ‘the offences and penalties in the Bill are there to demonstrate the seriousness with which we view the events that have taken place in the past and to act as a deterrent to any future misuse of human tissue’, (Human Tissue Bill, HL Deb 25 October 2004 vol 665 c1094).

To date the one person charged with a crime (under Section 32 of the Act) was a debt-ridden salesman, Dan Tuck, who attempted to sell his own kidney on the internet (http://www.thelawpages.com/court-cases/Daniel-Tuck-540-1.law). He would still have been found guilty of a crime had he been charged under the repealed Human Organ Transplants Act 1989 (the 1989 Act), but the sentence he received would have been quite different. At Wolverhampton Crown Court he was sentenced to a 12 month suspended jail sentence, a fine and 120 hours of community service. Had he been charged under the 1989 Act, he could not have been sentenced to more than a three month jail sentence or a fine (Section 1(5) Human Organ Transplants Act 1989). Controversially it could be argued that ‘respect for autonomy supports an individual who wishes to sell their organs or other human material’, (The Human Tissue Act 2004: Innovative Legislation – Fundamentally Flawed Or Missed Opportunity? page 181). That said, Section 32 of the Act is in line with Article 21 of the Convention on Human Rights and Biomedicine, and thus the argument is beyond the scope of this essay.
Where the Act provides criminal sanctions for non-compliance, civil law remedies are not provided and remain available only at common law. 'It is possible that compensation for wrongful retention or use in research might now be obtained through criminal injury compensation schemes. But otherwise wronged families and individuals will continue to find it difficult to obtain compensation', (Beyond Bristol And Alder Hey: The Future Regulation Of Human Tissue). The difficulty for the claimant arises in attempting to frame a claim in negligence by proving they have suffered a legally recognisable psychiatric injury – necessary as neither the Act nor the common law recognises that the misuse of or damage to human tissue which has been removed from the body constitutes a battery.

In Yearworth v North Bristol NHS Trust Court of Appeal (Civil Division) [2010] Q.B. 1 cancer patients’ frozen semen was thawed, rendering them unable to have children which, they claimed, led them to suffer nervous shock as a consequence. It was held that the semen was in fact ‘property’ and its damage was capable of establishing a claim in negligence without the need to prove they suffered psychiatric injury. The patients’ right to use the semen was limited by the Human Fertilisation and Embryology Act 1990, but this did not derogate from their ownership of it. Yearworth is a landmark case because until this decision the law did not recognise any proprietary interests in excised tissue. The case may be of limited application but could there be scope for development here?

In the context of the Act, relevant material intended to be used for transplantation is recognised as property 'because of an application of human skill' (Section 32(9)(c)). Why should this label not be applied to all human tissue removed from the body? In R. v Kelly (Anthony Noel) Court of Appeal (Criminal Division), [1999] Q.B. 621 an artist used the help of a morgue attendant to obtain body parts kept by the Royal College of Surgeons. It was held that property rights were created in those body parts through the application of skill and Rose LJ indicated that the application of skill may not even be necessary to create property rights. So why was this principle limited in the Act to transplantation material which has been subject to skill? 'The undefined nature of the work or
skill exception in the HTA 2004 also raises a secondary issue of what standard of skill should be required. It may be, for example, that allowing a drop of blood to fall onto a piece of filter paper for the purposes of analysis could be sufficient’, (Law and the Human Body: Property Rights, Ownership and Control p115).

If the skill exception were broadened or, as in Germany’s civil law jurisdiction, property rights in separated biological materials were recognised, it would give individuals’ a proprietary interest in their separated material which could provide a civil action where the material is misused, without the need to prove psychiatric injury. ‘Despite the instinctive revulsion towards the application of the concept of property to the human body and body parts, it has the benefits of predictability and accountability. Mason and Laurie have added that ‘recognising property rights in our person also facilitates further and better respect for individual autonomy, as is required by the ethical principle of respect for persons’, (Donated Organs, Property Rights and the Remedial Quagmire, Remigius N. Nwabueze, Medical Law Review 2008 16 (201)). A proprietary interest would provide a useful civil remedy yet the Act failed, bar the relevant material for transplantation exception, to change the common law ‘no property’ stance or to provide an alternative avenue for civil claims. A person may see a rogue doctor imprisoned or fined but they themselves are unlikely to succeed in claiming damages for the doctor’s actions; hardly in keeping with Dr Ladyman’s philosophy that the individual’s ‘body is theirs to control’ if they can find no statutory or common law redress for its misuse.

**Conclusion**

At 113 pages long, including 20 pages of explanatory notes, the Act is certainly unwieldy. The decision not to publish a draft bill could quite feasibly be at the root of the problem, but it was during the bill’s passage through Parliament that it developed ‘Frankenstein’s monster’ qualities, as sections were removed or added and revisions made. This has led to a piece of legislation in which Sections will routinely refer to multiple other Sections, so that the reader must refer forward or back numerous times on each page, in order to grasp the meaning they seek. As such it is poorly organised and does not readily assist those who
must comply with the law to function efficiently using the Act alone. The HTA’s 9 Codes of Practice and twice annual Legal Directions seek to clarify the law and assist practitioners in their compliance, yet as discussed earlier in this essay, there is a general ‘wooliness’ to many of the HTA’s instructions which could lead, quite lawfully, to a stark difference in standards of communication and practice across establishments.

The driving force of the Act was a move away from the 1961 Act’s assumption of permission to extract, store and use human tissue where there is ‘no reason to believe there is objection’, to a system of obtaining express ‘consent’ to do so. It is therefore startling that the Act neither fully defines ‘appropriate consent’ nor does it demand that consent be given in all circumstances; ‘there is a danger that the amendments made have led to a shift away from respect for individual autonomy and swung too far in favour of respect for the researcher’, (The Human Tissue Act 2004: Innovative Legislation – Fundamentally Flawed Or Missed Opportunity? page 186).

Further, it is a fiction to assume that all individuals would unanimously give consent to their tissue being used in the thousands of research projects granted ethical approval each year, yet the Act permits the tissue of living persons to be used in this way without their knowledge or approval. At the other extreme, consent and only consent will allow the analysis of a person’s DNA, when this could provide crucial familial information; can it truly be said that that information belongs solely to the individual when it is genetically shared? And directed organ donation is permitted in life, but after death a person may only give a generic consent to their organs being used by the community and any specific wishes will be ignored. The inclusion of criminal sanctions in the Act will certainly act as a deterrent to poor practice, but at present only one person has incurred a statutory penalty and he was from outside the medical community. And without the recognition of excised tissue as ‘property’, there is only the faintest glimmer of hope for the victim in terms of procuring a civil remedy via the common law.
The HRA has enshrined respect for the individual in UK law and the Act purports to be HRA compliant, yet the contentious issues discussed in this essay highlight the real possibility that the Act could be challenged under Articles 8 (right to respect for private and family life) and 9 (right to freedom of thought, conscience and religion) of the ECHR. As Gage J acknowledged in A v Leeds Teaching Hospitals NHS Trust Queen’s Bench Division [2005] Q.B. 506, a case brought under the 1961 Act by parents affected by the Alder Hey body parts retention; though each case will turn on its own facts, the extraction, storage and use of human tissue is capable of engaging Article 8 rights, without justification under Article 8(2). In A v Leeds the body part in question was a child’s brain, but there is no reason why the extraction, storage and use of other human tissue, from either living persons or cadavers, could not give rise to an infringement of a person’s ECHR rights.

‘Unwisely the government promised an Act that would create a clear and certain framework of rules. But as one would expect in this difficult ethical terrain, the product it delivered extends rather than diminishes the tangled web of statutes and common law principles that preceded it’, (Beyond Bristol And Alder Hey: The Future Regulation Of Human Tissue). In the House of Commons, Mr Paul Burstow claimed, ‘The Bill will put right the mistake of over 40 years ago by making it clear that consent must be obtained’, (Human Tissue Bill, HC Deb 15 January 2004 vol 416 c1007), but the Act has not fulfilled this prophecy. Parliament undertook a moral balancing act in its construction, weighing the competing biomedical business, research, community, familial and individual interests in the scales of the law. At its inception the rights of the individual weighed heavy in the scales, but by the time it gained royal assent, those rights were no longer the Act’s ballast. And as medical science continues to progress and society’s ethical expectations develop alongside, there will come a time, as with its predecessor, when a spotlight is shone upon the Act and its own inadequacies, inconsistent with respect for the individual, will be revealed.

FIONA SHORE
Bibliography

Brazier, M; Organ retention and return: problems of consent; J.Med Ethics 2003

Cronin, Antonia J; Directed and Conditional Deceased Donor Organ Donations: Laws and Misconceptions; Medical Law Review 2010

Hardcastle, Rohan; Law and the Human Body: Property Rights, Ownership and Control; Hart Publishing 2009

Herring, Jonathan; My Body, Your Body, Our Bodies; Medical Law Review 2007

Jones, D Gareth and Whitaker, Maja I; Speaking for the Dead: the Human Body in Biology and Medicine; Ashgate Publishing 2009

Liddell, Kathleen and Hall, Alison; Beyond Bristol And Alder Hey: The Future Regulation Of Human Tissue; Medical Law Review 2005

McGuinness, Sheelagh; Respecting the Living Means Respecting the Dead too; Oxford Journal of Legal Studies 2008


Nwabueze, Remigius N; Donated Organs, Property Rights and the Remedial Quagmire; Medical Law Review 2008

Price, David; The Human Tissue Act 2004; The Modern Law Review 2005

Taylor, Samantha and Wilson, David; The Human Tissue Act (2004), anatomical examination and the importance of body donation in Northern Ireland; Ulster Med J 2007
**ADR, UK Sport and the structure of Tribunals**

DHARMINDER RUPRA

With the increase in sports disputes, it is arguable for sports tribunals to be given a statutory footing within the tribunal structure.

The Tribunal, Courts and Enforcement Act 2007 (TCEA) established a First Tier Tribunal and an Upper Tribunal. The TCEA provides the power and appeal rights to and from these tribunals.\(^1\) The TCEA provides a constitutional settlement for tribunals and their judiciary. This new ‘tribunal system’ has brought some tribunals into its structure, however the rest remain outside.

Sport Resolutions UK (SRUK) is an independent and private dispute resolution service for sport in the UK. The Arbitration Act 1996 (AA) provides the legal framework within which domestic arbitration operates. The English courts’ approach to sports-related disputes is that they are better dealt with ‘in-house’ by way of disciplinary mechanisms using arbitration. ‘While every court is a tribunal, the converse is not true’.\(^2\)

Arbitration is the procedure where a dispute is submitted by agreement of both parties to an arbitrator who will then make a binding decision on the dispute: this comes under the remit of Alternative Dispute Resolution (ADR). Sport provides a good example of the benefits of ADR in comparison to judicial proceedings. The TCEA is consistent with this approach.\(^3\) The main advantage of ADR over litigation is that the parties voluntarily agree to enter into the ADR process and be bound by a decision of an independent adjudicator of their choice [party autonomy]. Other advantages which coincide with the TCEA include cost-effectiveness and speed,\(^4\) as well as the ability to use adjudicators who are

---

\(^1\) TCEA s.11(3) and s.13(3).
\(^3\) TCEA s.2(3)(d).
\(^4\) TCEA s.2(3)(b)(ii).
experts in the field of dispute. The principles of procedure give the Senior President its duties. There also exists the objectives for the rules of procedure.

SRUK offers arbitration services for sports bodies in the UK where hearings are held in private (decisions are not usually published unless both parties agree to them being so). This is an important factor for some sports in the UK which receive huge amounts of media attention. The AA provides the governing rules and codes as to the arbitration service which SRUK offer. SRUK’s ‘tribunal appointment’ service is, however, the most popular of form of resolving disputes. This is a service whereby high performance professionals or community sports bodies request that a member of one of SRUK’s specialised list of arbitrators be designated to chair and resolve an internal dispute. The advantage of this procedure for sports bodies is that instead of having to create and maintain their own independent appeals committee, they can on an ad hoc basis call upon the expertise of SRUK’s panellists (now used by leading sports organisations as part of their ADR mechanism), who offer a broad level of experience and specialisation across the range of sports dispute-related areas.

Common areas of concern include discipline, anti-doping, child welfare, personal injury, intellectual property, commercial and employment law, and professional negligence. This wide range of sports-related disputes and sports bodies reflect SRUK’s growing importance in the modern administration and regulation of sport in the UK. The tribunal appointment procedure ties in with SRUK’s main objective, which is ‘to make available to all sports in the UK, independent, expert, timely and cost-effective resolution of all disputes’.

Essentially SRUK’s increasing authority and competency is owed to the fact that it is the provider of the National Anti-Doping Panel (NADP), which is the independent body established to determine anti-doping disputes within the UK. Its main purpose is to improve the quality and consistency of tribunal decision-

---

5 TCEA s.2(3)(c).  
6 TCEA s.2.  
7 TCEA s.22(4).  
8 www.sportresolutions.co.uk
making in anti-doping cases. SRUK should receive formal governmental recognition of its role and come within the tribunal system in the UK, Administrative Justice and Tribunals Council and the TCEA.

If SRUK received greater acknowledgement and came within the tribunal system such a body would develop a set of decisions and awards that would be of useful precedential value in the resolution of sports disputes. ‘It will be principally for the [new] upper tribunal to lay down guidelines as to the precedent effect of its decisions for different purposes’.9 This would ensure equal treatment of aggrieved sports parties if similar situations arose.

Sports disputes are also becoming increasingly complex in content and legalistic in form and therefore even more difficult to deal with for sports bodies under their internal structures and within their own resources. Commercial issues are now at the forefront of sporting disputes more than ever before and there is no better example of these disputes than can be seen in professional football.

In recent years an unprecedented number of professional football clubs have found themselves in severe financial difficulties, eventually entering administration. The Football Association has a points sanction where ‘if a club enters administration during a league season 15 points will be deducted from the points total of that club’,10 usually resulting in the relegation of the club from the league. This was seen recently in the 2009/10 Premier Football League where Portsmouth Football Club entered administration, were ‘docked’ 15 points and eventually relegated to the Championship league. The loss in revenue was immense and cost the club an amount in the region of £40m through the loss of ‘live’ televised games revenue and sponsorship.

Undoubtedly the most controversial saga illustrating these ever-increasing commercial issues in sports disputes is the ‘Tevez’ saga.11 The footballer Carlos

---

9 Cadogan v Sportelli [2008] 1 WLR 2142 at [99] as per Carnwath LJ.
10 Football League’s (Insolvency) Policy.
11 ‘The Mysterious Case of Who Owns Carlos Tevez’ – P.Kelso The Daily Telegraph, Sport Soccer 15/05/09.
Tevez joined West Ham United on loan in August 2006. It was later revealed that an off-shore company owned his registration and accompanying economic and commercial rights. This arrangement was in breach of the Football Association Premier League (FAPL) rules relating to ‘third party’ ownership of players. These rules were designed to minimise the influence that outside entities might have on the operation of the Premier League. The FAPL thus set up an independent disciplinary tribunal to investigate the continuing West Ham/Tevez arrangement.

At the tribunal West Ham pleaded guilty to breaching FAPL rules and were fined £5.5M. West Ham agreed to end the contractual arrangements with the third party owners and controversially were not ‘docked’ any league points for breaching FAPL rules. This was significant as the hearing took place alongside West Ham’s relegation battle to remain in the Premier League. A points deduction would probably have seen them demoted to the Championship, causing a loss of up to £40m in TV-related revenue. Eventually Tevez contributed enormously to West Ham’s victories in the 3 final games of the season and helped them from an unlikely escape from relegation.

Sheffield United were however relegated and, arguing that West Ham’s breach of FAPL rules had been the effective cause, pursued the matter to independent arbitration and in court. In March 2009 the parties reached an out of court settlement where West Ham agreed to pay an estimated £20M in compensation to Sheffield United.

The ‘Tevez’ saga resulted in a complex array of disciplinary, arbitral and court hearings. If such disputes are a sign of what is to come in the future then only a statutory sports tribunal in the UK would have the capacity and competence to

---

12 ‘No Justice if Tevez Keeps West Ham Up’ – D.Bond The Daily Telegraph, Sports Soccer 28/04/07.
14 The High Court challenge was heard in private.
15 ‘West Ham to pay £20M over Tevez Claim’ – P.Kelso The Daily Telegraph, Sport Soccer 17/03/09.
16 Sheffield Utd FC Ltd v West Ham Plc. [2008] EWHC 2855 (Comm.).
deal with them. This sports tribunal would assist sports bodies in resolving disputes that may sometimes have complex legal arguments without burdening the courts with proceedings that remain at core sport-specific in substance. Certain features that distinguish a tribunal from a court are recognised in statute.17

A body such as SRUK would have the necessary expertise and independence to handle sporting disputes effectively. This would remove the financial and administrative burden of such cases for football authorities and eliminate any possible conflicts of interest and improve the quality and consistency of decision-making if similar cases were to arise. The time has now come to give de jure recognition to SRUK’s competency and authority. In the interest of clarity, effectiveness and consistency, SRUK should be placed on a statutory basis under the TCEA. There are other countries that have adopted this approach, e.g. the Sport Dispute Resolution Centre of Canada (SDRCC)18 and the Sports Tribunal of New Zealand (STNZ).19 Similar to tribunals under the TCEA, they follow a judicial process but proceedings are conducted in an informal manner.

This proposal of a new system to deal with sports disputes in the UK would involve the initialisation of a new Sports Tribunal in the UK. This new system would help to deal with the increase in sporting issues with which SRUK has to deal at present. It is the less well-resourced sports bodies that mainly seek SRUK’s services. The better-resourced sports bodies such as the Football Association and the Jockey Club have the advantage of using their own independent arbitral mechanisms. However, all of these arbitral services are bound by adherence to private codes of conduct and the AA.

The Court of Arbitration for Sport (CAS) based in Lausanne, Switzerland has gradually built up the trust of the sporting world such that it has earned its status as one of the ‘principal mainstays’ of the administration of international sport. Recently tennis player David Savic failed in his appeal against a life ban for

---

17 TCEA s.2(3).
18 www.crdsc-sdrcc.ca
19 www.sporttribunal.org.nz
match fixing. CAS provided confirmation of the ruling reached by the Tennis Integrity Unit. He is only the second player to receive a life ban following Daniel Koellner who was also found guilty of match fixing in May 2012.

In recent years there has been a significant surge in the number of women entering professional sports. In November 2006 the Luton Town FC Manager Mike Newell launched a scathing attack on women officials. Assistant referee Amy Rayner angered Mike Newell as she failed to award his side a penalty in their 3-2 loss to QPR. He said that ‘bringing women into the game of football is not the way to improve refereeing’.20 In Rayner’s defence the FA’s Head of Senior Referee Development, Neale Barry, said ‘Amy is one of our leading female referees, English football needs more Amy Rayners’.21 Similarly ‘Sky Sports’ commentators Andy Gray and Richard Keys’ sexist comments on 22nd January 2011 about Sian Massey, a female football official, sparked controversy. They said that ‘she and other female assistant referees did not know the offside rule’. Keys said ‘somebody better get down there and explain offside to her’, and Gray replied ‘women don’t know the offside rule’. Gray was subsequently sacked from his position for ‘unacceptable and offensive behaviour’ and Keys resigned, stating that his position was ‘untenable’.22

The issue of sexism does not, of course, exist only in football. Indeed there are many different sports where women are actively involved but do not receive the recognition they deserve. To prevent these issues arising frequently a statutory Sports Tribunal in the UK would serve to provide a solid foundation for their resolution. This type of tribunal would come under the general supervision of the Administrative Justice and Tribunals Council which under the TCEA replaces the Council on Tribunals.

With an increasing level of finances entering the sporting arena the analysis of controversial issues are under extensive coverage by the media. In fact there have been circumstances where reporters have been purporting to be

20 www.bbc.co.uk/sport/rayner
21 www.bbc.co.uk/sport/rayner
22 www.bbc.co.uk/sport/massey
businessmen and arranging meetings with high profile sports players. These reporters have enticed sports players with lucrative offers to take part in some form of disrepute, either by asking them to deliberately lose games or arrange certain events to happen during the course of a sporting event. The reason for this is the large sums of money that surround betting and sport. Certain parties can make substantial amounts of money through this illegal practice.

In 2010 World Champion snooker player John Higgins had purported to lose certain matches for a financial reward. The issue once brought to light was heard by the independent SRUK tribunal, which is named within the World Professional Billiard and Snooker Association (WPBSA) disciplinary rules. Higgins was fined £75,000 and he was banned from snooker for 6 months.23

A sports tribunal given formal governmental recognition under the TCEA will be able to deal with such contentious issues by using its authority in an effective and fair way to reduce the likelihood of these events occurring. At present SRUK’s service is bound by the AA 1996 and there is now a need for this service to come within the remit of the TCEA 2007. SRUK would then fit into the rules of procedure24 and practice directions,25 acting judicially26 and the Senior President must adhere to his/her duties27 by making tribunals, amongst other requirements, accessible28 and fair.29

If the UK government decides to adopt a similar approach to the tribunals in New Zealand and Canada there would then exist a greater system of uniform consistency to tackle the complex disputes that so often arise in sports. The aim is for tribunal decisions to be reached fairly, and to be well reasoned, speedy and timely. The tribunal should give a written decision and an explanation of the

23 www.sportresolutions.co.uk/news/higgins
24 TCEA s.22.
25 TCEA s.23.
26 TCEA s.1.
27 TCEA s.2(3).
28 TCEA s.2(3)(a).
29 TCEA s.2(3)(b)(i).
reasons for the decision to all parties.\textsuperscript{30} This would serve to provide confidence in the system being independent, open, fair, accessible, cost-effective and impartial.

We have to be aware that the UK is highly prominent in the sporting world. This has resulted in an increase of people taking to sports and international sporting events such as the 2012 Olympics\textsuperscript{31} and the 2015 Rugby World Cup,\textsuperscript{32} which are both being held in London. This is likely to encourage more members of society to join sports clubs and to participate in sport.

Some sporting disputes are minor and may be simply resolved without recourse to formal procedures. However, as this article has argued, the more complex issues regarding professional sports will become easier to resolve only if SRUK is brought within the remit of the TCEA. This will give SRUK a more responsible role, and recognising that it is able to provide a resolution to sporting disputes is a positive step forward for sports in the UK.

DHARMINDER RUPRA

\textsuperscript{30} Scott v Scott [1913] AC 417.
\textsuperscript{31} www.london2012.com
\textsuperscript{32} www.rugbyworldcup.com
I often receive a shocked response when I tell those I meet that I am studying to become a barrister. Whether this is due to the fact that I am blonde, female, speak with a Yorkshire accent or an amalgamation of all three is a mystery. My accent, however, has been the characteristic that has received the most comment throughout my legal learning. The first question those I meet tend to ask is ‘where is that accent from?’ (I even met one gentleman who innocently asked if I lived near a coalmine!) I am now due to embark on the long anticipated application for pupillage and I can only wonder if this stereotyping will continue and if it indeed exists in the legal profession today. Is there a place for a regional accent in the courtroom, or is the Bar reserved for the silver-tongued Queen’s English speaking advocate?

Prior to leaving my small town of Yarm in North Yorkshire, I was convinced that I simply did not have an accent. Surrounded by North Yorkshire speaking folk up to the age of eighteen, it wasn’t until I left the comforts of home to commence my undergraduate Law degree that I became aware that I didn’t possess the traditional BBC Newsreader dialect that most of my fellow law students seemed to pull off so successfully. I remember an enthusiastic law student from Kensington approach me on registration day and ask:

‘What do you read?’
‘Law’ I replied.
‘Art?’
‘No, Law...’
‘Art?’

Whether this was due to inexperience in understanding an accent or the result of a stereotype, I will never know. This conversation, believe it or not, occurred whilst I was studying at Leeds University.

It appears, however, that the media is welcoming the move away from the stereotypical barrister. The BBC’s recent portrayal of a buzzing London
chambers can be seen in the drama series Silk, whose heroine ‘Martha Costello’ is a tough-talking barrister with a thick Bolton accent. However, actress Maxine Peake who plays the role comments that she was asked to soften her own rich Bolton accent for the series:

"For Silk I had to soften my accent. They go, 'Okay, so this character is from the north but she went to university, Maxine, and has lived in London for 10 years.' So I went, 'OK, well I went to Rada and lived in London for 13 years,' and they go, 'Yeah, but she’s lost her accent a bit more than you have,'”

– implying that Barristers are perceived as having a thinner Bolton accent than actresses.

However, there is hope for my fellow flat-vowelled advocates. A criminal barrister practising in London sees ‘having an accent’ as an advantage. He feels that his Northern Irish accent allows him to stand out from the crowd and relate more to the jury. However, he states that how his accent is received depends on the audience listening:

"You can go into a Divisional Court and have Judges react in a different way than a jury can”.

He feels that in particular circumstances it can pose a disadvantage such as in front of certain Judges. He feels that barristers who are discriminated against more frequently are those with a non-English accent. He once overheard counsel making an accented barrister the butt of their jokes whilst court had adjourned. He has also experienced some comments about his accent in the past, similarly in jest, when he was asked if he was a refugee from ‘the troubles’. A student from Essex also once informed him that she was told to lose her accent during advocacy training, encouraging her to conform to the barrister stereotype.

It seems that vocal training is becoming more popular with leading law firms, as they send their advocates on training courses to hone their vocal and public

---

speaking skills. Is this to target certain advocates with accents, or to improve their general advocacy skills? I asked Luan De Burgh, Presentation and Communication Coach of The De Burgh Group if he has ever coached anyone in the legal profession.

"A great many lawyers have come to us for presentation and communication skills training and that has included a number of barristers. I have found when working with barristers that it is often more a case of speech clarity rather than broader accent that is of importance".

Mr De Burgh does however feel that as soon as somebody opens their mouth to speak they are judged by their accent.

"Like it or not, there are preconceived stereotypes that accompany certain accents and hearing those accents can ignite prejudices."

Mr De Burgh agrees that barristers are stereotypically seen as having an ‘RP’ accent by a large proportion of the population; RP being shorthand for Received Pronunciation, or in other words ‘the Queen’s English’.

"This comes from an outdated stereotype of the public school, Oxbridge barrister. In 2012, there are many more barristers who have not come from either of those two routes and who, collectively, speak with any number of British accents and this is becoming more the norm."

It appears that with my short experience at the Bar thus far, I have encountered a diversity of accents. But has this eclectic mix of accents quashed the ‘outdated’ stereotype, or does prejudice still exist in the legal profession?

Rachel Donald, a family solicitor practicing in Newcastle explains how her North Eastern accent has affected her career:

“I have received comments about having a regional dialect, but usually not in a derogatory sense. It depends upon whether the client also has a regional accent but I think that clients generally find you more approachable.”

Miss Donald states that even in Newcastle, judges with a regional accent are in the minority. She also highlights the advantage of having an accent:
“Clients always seem to relate to your comments more than the Judges who speak in RP.”

Miss Donald admits that she adopts a more ‘neutral’ accent when addressing the Court, perhaps as a way to conform to the traditional RP accent of Court. Mr De Burgh finds generally that lawyers do adapt their accent as a way to ‘fit in’, but speaks of “those individuals who stand out by not conforming and being of the ‘this is who I am – take it or leave it’ group”.

I spoke to a solicitor from Manchester who seemed to fit this description. She is confident that her Mancunian accent has had no effect on her legal career in London. Although as a child she received elocution lessons, she feels that any discrimination those with accents feel is possibly all in the head of the recipient. Similar to the barrister mentioned earlier, she promotes the advantage of being different and standing out from the crowd in a highly competitive profession.

Whilst attending an advocacy training session it became apparent that I myself have conformed to the RP barrister stereotype without even being aware of the transformation process. Whilst the judge commended a student (to her delight) for her Welsh accent, he proceeded to list the advantages that an accent brings. It helps dispel the traditional layman’s view that all barristers are pompous, public school boys that do not give a second thought to Joe Public. He explained his view that an accent encompasses the ‘likability factor’ and is less intimidating to clients. In addition to the beaming Welsh student he sent encouraging advice in the direction of a Scottish student. I eagerly waited in my seat with my flag ready to fly for North Yorkshire but waited in vain as the judge continued the training without a word of praise for my accent. It wasn’t until I spoke to those in my group after the exercise and received the response ‘what accent?’ that it dawned on me. Without even noticing, as Maxine Peake adopted a more neutral Northern accent for the BBC’s portrayal of the twenty-first century barrister, it seems that I have toned down my accent to address the court.

Although I have been told that perhaps the prejudice I experience is ‘all in my head’, I find this hard to accept. However, in my experience it seems that
constant mimicking and remarks of jest have gradually waned, having occurred mostly at University. I can only assume that this is because students are sheltered youths straight from their home town and haven’t so much as spoken to an individual North of the M25! Because of this they are led to presume the well-dated stereotype that every northerner is poverty stricken but still manages to afford a flat cap. After speaking with professionals in the legal field, it seems that students quickly learn this is not the case and hopefully life experience will iron out these prejudices.

It seems that there is hope for the non-RP speaking advocate. Those I have spoken to promote having an accent as a positive trait that clients find less intimidating and presents a flare of individuality. However in my view, albeit in a positive way, this is still a form of judging a book by its accent. In a highly competitive, predominantly self-employed profession, there is a constant need to stand out from the crowd. Although I will not be adding ‘Northern Accent’ to my Curriculum Vitae, perhaps at interview for Pupillage my accent will show a sense of individuality and work in my favour.

CHARLOTTE DONALD