

# KING'S STUDENT LAW REVIEW

# KING'S STUDENT LAW REVIEW

2(1) 2010

## EDITORIAL BOARD

EDITOR-IN-CHIEF

Hin-Yan Liu

MANAGING EDITOR

Olivia Wybraniec

SENIOR EDITOR

K Mark Sun

EDITORS

Toon Hylebos

Siubhan Magee

Matthew Smith

Thomas MacManus

Sophie Murray

Riddhi Shah

SUBMISSIONS EDITOR

Jonny Bushby

WEB EDITOR

Laure Kroely

FACULTY ADVISOR

Prof. Penny Green

<http://www.kslr.org.uk>  
[editors@kslr.org.uk](mailto:editors@kslr.org.uk)

The King's Student Law Review seeks to publish the best of student legal scholarship. The KSLR is generously supported by the King's College London Law School and the King's College London Annual Fund.

Published in the United Kingdom  
By the King's Student Law Review  
King's College London School of Law  
Strand London  
WC2R 2LS

In affiliation with the King's College London School of Law  
<http://www.kcl.ac.uk/law>

All rights reserved.

No part of this publication may be reproduced, transmitted, in any form or by any means, electronic, mechanical, recording or otherwise, or stored in any retrieval system of any nature, without the prior, express written permission of the King's Student Law Review.

Within the UK, exceptions are allowed in respect of any fair dealing for the purpose of research of private study, or criticism or review, as permitted under the Copyrights, Designs and Patents Act, 1988. Enquiries concerning reproducing outside these terms and in other countries should be sent to the Editor-in-Chief of the KSLR at the above address.

© King's Student Law Review 2010

# TABLE OF CONTENTS

The EU Maintenance Regulation: A qualified success for European Family Law.....	5
---	---

Philip Bremner

Moral Hazard and How It Was Invoked in the Northern Rock Crisis of 2007.....	27
--	----

Andrés Curia Miranda

Why the Criminal Law Principle of Correspondence Should Prevail Over the Pragmatism of Constructed Liability.....	43
---	----

Dominic JJ Bright

Selective Abortion: Selecting the right response.....	55
---	----

Chris Cowland

Procreative Liberty and Selecting for Disability: Section 14(4) Human Fertilisation and Embryology Act 2008.....	71
--	----

Edward M Taylor

It's Just Not Cricket: EU Competition Law and its application to sporting rules.....	87
--	----

Eoin Kealy

The right to life shall be secured to everyone by law: The extent to which the European Court of Human Rights has developed the concept of positive obligations in relation to Article 2.....	103
---	-----

Louise McCamphill



# THE EU MAINTENANCE REGULATION: A QUALIFIED SUCCESS FOR EUROPEAN FAMILY LAW

PHILIP BREMNER  
UNIVERSITY OF ABERDEEN  
P.BREMNER.04@ABERDEEN.AC.UK

**ABSTRACT** The EU's Maintenance Regulation has recently been completed and the UK has finally decided to opt in. Although there are parallel developments on the international front, at The Hague Conference on Private International Law, this regulation appears to be a welcome addition to EU Private International Law. Despite the necessary reform instituted by the regulation, the question remains whether this has been achieved in the best possible way. In a union of 27 member states it is no small victory to achieve any sort of agreement on matters of Private International Law, which tend to be technical in nature and subject to a high degree of divergence between the member states. This, however, does not obviate any discussion of the legislative process leading to the regulation. It is important to ascertain whether the correct legislative procedure was used, as informed by the choice of legal basis. Far from being of mere academic interest, the competency issue goes to the very credibility and constitutionality of the legal instrument. In addition, an analysis of the UK's role in the legislative process can be highly instructive as to the way that, amongst others, Private International Law measures are being adopted at the EU level. The UK's use of its opt-in as a negotiating tactic is seen as an uncooperative and disingenuous way of advancing Britain's own interests, which could be achieved, less antagonistically, through ordinary legislative means. However, by not opting-in initially, the UK was able to take its international commitments at The Hague more seriously and, in the end, change the content of the regulation itself. Rather than an attempt to undermine the undoubted progress the EU has made in the advancement of Private International Law in Europe, this analysis, by discussing both the substance and the adoption of the regulation, is intended to reveal how this might be better achieved in the future. The conclusion is that it is not only implementing the desired reforms that is important, but also that the appropriate procedure is followed in their adoption and that the EU is mindful of its international obligations when adopting internal legislation.

## Introduction

The EU's latest foray into the realm of Private International Law, in the form of the recently adopted Maintenance Regulation,<sup>1</sup> comes close to achieving its ambitious objectives, but still leaves something to be desired. The ends, however, do not justify the means. Although the EU engaged in inter-institutional consultation and eventually played its role as a responsible international actor, it needs to take the legislative role of the European Parliament more seriously and defer to the more global Hague Conference on Private International Law more readily, especially where International Family Law is concerned.

The regulation builds on the Maintenance Convention<sup>2</sup> concluded relatively recently in The Hague, achieving closer regional cooperation within the EU than was possible at the more global Conference on Private International Law. The interaction between these two processes seems to have been well managed, playing to each other's strengths, and the resulting instruments are all the stronger for that. In order to understand fully the strengths and weaknesses of the Regulation, however, it is necessary to analyse both its substance and the means by which it was adopted. In terms of substance, this analysis reveals that the broad jurisdictional rules and abolition of *exequatur*, desirable as they are, come at a price, namely a two-speed Europe. As regards the adoption procedure, there is some controversy surrounding the choice of legal basis and the tactical use during the negotiations of the UK's ability to opt-in. By examining these issues, the present writer aims to illustrate how the Maintenance Regulation is a qualified success for both international cooperation and regional integration, but still leaves some lessons to be learned for the future.

## Background

The reform of the existing law relating to international maintenance obligations began at The Hague Conference of Private International Law. It was only after work had begun at The Hague that the EU decided to begin its own process of reform.<sup>3</sup> The work at The Hague was an effort to update and revise the pre-existing conventions relating to maintenance, in an attempt to facilitate its speedy recovery. Whilst the EU also aimed to achieve this; its primary goal was the abolition of *exequatur*<sup>4</sup> which, under Brussels I,<sup>5</sup> was required, inter alia, for decisions relating to maintenance claims.

There already exists a number of international<sup>6</sup> and European<sup>7</sup> instruments concerning international maintenance obligations and a new one has recently been concluded in The Hague. Given the number of existing measures in this area, the EU needs to make a compelling case for the necessity of further regulation. In the present writer's submission, the EU has made such a case.

<sup>1</sup> Council Regulation (EC) 4/2009 of 18 December 2008 on jurisdiction, applicable law, recognition and enforcement of decisions and cooperation in matters relating to maintenance obligations (the Maintenance Regulation) [2009] OJ L 7/1.

<sup>2</sup> Convention of 23 November 2007 on the international recovery of child support and other forms of family maintenance (the Maintenance Convention) <[http://www.hcch.net/index\\_en.php?act=conventions.text&cid=131](http://www.hcch.net/index_en.php?act=conventions.text&cid=131)> accessed 1 April 2009.

<sup>3</sup> P Beaumont, 'International Family Law in Europe – the Maintenance Project, the Hague Conference and the EC: A Triumph of Reverse Subsidiarity' (2009) 73 *Rabels Zeitschrift für ausländisches und internationales Privatrecht* (forthcoming).

<sup>4</sup> Council Draft programme of measures for implementation of the principle of mutual recognition of decisions in civil and commercial matters (30 November 2000) [2001] OJ C 12/1.

<sup>5</sup> Council Regulation (EC) 44/2001 of 22 December 2000 on jurisdiction and the recognition and enforcement of judgments in civil and commercial matters (Brussels I) [2001] OJ L 12/1.

One of the most compelling arguments supporting the need for the Maintenance Regulation is the simplification it will provide to the recovery of international maintenance obligations within the EU. Firstly, the regulation will take precedence<sup>8</sup> over the pre-existing Hague Maintenance Conventions and the New York Maintenance Convention.<sup>9</sup> This will avoid problems stemming from the fact that not all member states have ratified those conventions and various member states have made different reservations. In that regard, the Maintenance Regulation will level the playing field.

However, this alone might not be enough to demonstrate that the regulation is necessary. In addition, the Maintenance Regulation brings added value that only it could achieve. The Regulation represents an overhaul of the Brussels I regime, eliminating the remaining obstacles to the recovery of international maintenance obligations such as restrictive jurisdictional rules and intermediate enforcement measures. This was only possible because of the close regional integration within the EU. Such innovation could not have been achieved at the global level because of lack of consensus.

## Adoption Procedure

### Legal Basis

Since the Treaty of Amsterdam,<sup>10</sup> the EU has had express legislative competence to adopt “measures in the field of judicial cooperation in civil matters”<sup>11</sup> which have “cross border implications... insofar as necessary for the proper functioning of the internal market.”<sup>12</sup> This includes measures relating to jurisdiction, applicable law, recognition and enforcement.<sup>13</sup> The Maintenance Regulation was adopted on this basis in accordance with the procedure laid down in article 67(2) EC.

Since the Treaty of Nice,<sup>14</sup> Article 67(5) EC<sup>15</sup> has laid down two procedures for the adoption of measures in this area. It provides that all measures within the field of judicial cooperation in civil matters are to be adopted using the co-decision procedure,<sup>16</sup> “with the exception of aspects relating to family law,”<sup>17</sup> which requires unanimity in the council and mere consultation with the European Parliament. The European Commission urged the

<sup>6</sup> Convention on the law applicable to maintenance obligations towards children (adopted 24 October 1956, entered into force 1 January 1962) 510 UNTS 161; Convention on the law applicable to maintenance obligations (adopted 2 October 1972, entered into force 1 October 1977) 1056 UNTS 199; Convention of 15 April 1958 concerning the recognition and enforcement of decisions relating to maintenance obligations towards children (adopted 15 April 1958, entered into force 1 January 1962) 539 UNTS 27; Convention on the recognition and enforcement of decisions relating to maintenance obligations (adopted 2 October 1972, entered into force 1 August 1976) 1021 UNTS 209; Convention of 20 June 1956 on the recovery abroad of maintenance (adopted 20 June 1956, entered into force 25 May 1957) 268 UNTS 3 (New York Maintenance Convention).

<sup>7</sup> Brussels I Regulation (n 5); Regulation (EC) 805/2004 of the European Parliament and of the Council of 21 April 2004 creating a European Enforcement Order for uncontested claims (EEO) [2004] OJ L 143/15.

<sup>8</sup> Maintenance Regulation (n 1), Art 69.

<sup>9</sup> n 6.

<sup>10</sup> Treaty of Amsterdam amending the Treaty on European Union, treaties establishing the European Communities and certain related acts (Treaty of Amsterdam) [1997] OJ C 340/1.

<sup>11</sup> Treaty establishing the European Community (EC Treaty, as amended) [2002] OJ C 325/33, Art 61(c).

<sup>12</sup> EC Treaty (n 11), Art 65.

<sup>13</sup> *Ibid.*

<sup>14</sup> Treaty of Nice amending the Treaty on European Union, the treaties establishing the European Communities and certain related acts (Treaty of Nice) [2001] OJ C 80/1.

<sup>15</sup> “By derogation from paragraph 1, the Council shall adopt, in accordance with the procedure referred to in article 251... the measures provided for in Article 65 with the exception of aspects relating to family law.”

<sup>16</sup> Article 251 EC, which allows the European Parliament detailed involvement in the drafting of the regulation and requires agreement of both the Council and the European Parliament.

<sup>17</sup> n 15.



Council to use the Passerelle procedure in Article 67(2) EC,<sup>18</sup> to bring maintenance obligations within the co-decision procedure and to require only qualified majority voting in the Council.<sup>19</sup>

The draft proposal for such a decision, which the Commission communicated to the Council, was supported by convincing arguments for the use of the co-decision procedure.<sup>20</sup> The Commission stressed that, although undoubtedly connected with family law, maintenance obligations are primarily civil debts and are thus mainly a financial matter:

... it is important to properly reflect the hybrid nature of the concept of maintenance obligation – a family matter in origin but a pecuniary issue in its implementation, like any other claim.<sup>21</sup>

This reasoning is supported, as the Commission notes,<sup>22</sup> with reference to Brussels I, which excludes family law matters but still deals with maintenance obligations.<sup>23</sup> This stands in contrast with Brussels II bis, which deals with core family law matters but leaves maintenance obligations out of account.<sup>24</sup> Furthermore, the Commission's assertion "that recovery of a maintenance claim does not go to the core of [family] relationships"<sup>25</sup> is supported by the fact that the Maintenance Regulation provides that "the recognition and enforcement of a decision on maintenance... shall not in any way imply the recognition of the family relationship, parentage, marriage or affinity underlying the maintenance obligation which gave rise to the decision."<sup>26</sup>

Despite the Commission's compelling arguments, it chose not to present a formal proposal largely because the majority of member states would not support it and therefore the unanimous agreement in the Council that would be required could not be achieved.<sup>27</sup> This is regrettable from the point of view of the democratic legitimacy of the Maintenance Regulation. The European Parliament clearly sided with the Commission on this one,<sup>28</sup> agreeing that:

Once an obligation to pay maintenance has been established under family law, what we are left with is simply a pecuniary obligation – a debt like any other... a

<sup>18</sup> "...the Council, acting unanimously after consulting the European Parliament, shall take a decision with a view to providing for all or parts of the area covered by this Title to be governed by the procedure referred to in Article 251..."

<sup>19</sup> Commission of the European Communities, 'Communication from Commission to Council' (Brussels 15.12.2005) COM (2005) 648 final <<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2005:0648:FIN:EN:PDF>> accessed 1 April 2009.

<sup>20</sup> *Ibid.*

<sup>21</sup> European Commission (n 19) 4; See also Letter from Rt Hon Baroness Ashton of Upholland, Parliamentary Under Secretary of State, Department for Constitutional Affairs to the Chairman of European Union Select Committee, European Union Select Committee, 'Correspondence with Ministers' HL (2006-2007) 187, 380 <<http://www.publications.parliament.uk/pa/ld200607/ldselect/ldcom/187/187.pdf>> accessed 1 April 2009.

<sup>22</sup> European Commission (n 19) 4.

<sup>23</sup> Brussels I Regulation (n 5), Art 5(2).

<sup>24</sup> Council Regulation (EC) 2201/2003 of 27 November 2003 concerning jurisdiction and the recognition and enforcement of judgments in matrimonial matters and the matters of parental responsibility, repealing Regulation (EC) No 1347/2000 (Brussels II bis) [2003] OJ L 338/1, Art 1(3)(e).

<sup>25</sup> Commission of the European Communities (n 19) 5.

<sup>26</sup> Maintenance Regulation (n 1) 22.

<sup>27</sup> See Baroness Ashton (n 21) 380; M Hellner, 'The Maintenance Regulation: A Critical Assessment of the Commission's proposal' in K Boele-Woelki and T Sverdrup (eds), *European Challenges in Contemporary Family Law* (European Family Law, Intersentia, Antwerp 2008) 347.

<sup>28</sup> European Parliament legislative resolution of 13 December 2007 on the proposal for a Council regulation on jurisdiction, applicable law, recognition and enforcement of decisions and cooperation in matters relating to maintenance obligations (COM(2005)0649 – C6-0079/2006 – 2005/0259(CNS)) [2008] OJ C 323 E/470 Amendment 3.

maintenance obligation is a pecuniary claim and the fact that it arose out of a family or similar relationship has scarcely any relevance.<sup>29</sup>

Just as the European Commission did, the European Parliament supported its position with reference to other instruments dealing with maintenance. In this regard, the Committee on Legal Affairs on the Proposed Legal Basis cited the European Enforcement Order Regulation,<sup>30</sup> which both deals with maintenance and was adopted using the co-decision procedure.<sup>31</sup>

The Commission, however, notes that the difference between Brussels I and the EEO, on the one hand, and the Maintenance Regulation, on the other, is that maintenance obligations are only dealt with in the former as an ancillary matter, whereas they are the sole concern of the latter. This difference necessitates, as a matter of law, that a different procedure is used. However, this does not negate the underlying claim that the regulation does not affect the “fundamental nature and expression of personal relationships between members of the family” and therefore does not deserve the special treatment that the rest of family law receives.<sup>32</sup>

### ***UK's Opt-in***

The default position, following the Protocol<sup>33</sup> annexed to the EC<sup>34</sup> and EU<sup>35</sup> Treaties by the Treaty of Amsterdam,<sup>36</sup> is that the UK does not participate in instruments adopted on the basis of Title IV<sup>37</sup> of the EC Treaty unless it opts in. This can be done either within three months of the Commission's proposal<sup>38</sup> or at any time after the adoption of the instrument.<sup>39</sup> In the case of the Maintenance Regulation, the UK chose not to opt in before the instrument's adoption because of concerns about the applicable law provisions in the Commission's proposal, and also because of a desire not to compromise its negotiating position in The Hague.<sup>40</sup> However, once the negotiations in The Hague had been concluded and a compromise over applicable law had been reached, the UK declared its intention to opt in on 28 November 2008.<sup>41</sup>

The way in which the UK used its opt-in to shape the content of the Maintenance Regulation raises the question of its desirability as a negotiating tactic. The privileged position of the UK and Ireland as regards Title IV EC was a necessary concession made by the other EU member states during negotiations over the Treaty of Amsterdam in order to

<sup>29</sup> Opinion of the Committee on Legal Affairs on the Proposed Legal Basis, ‘Report on the proposal for a Council regulation on jurisdiction, applicable law, recognition and enforcement of decisions and cooperation in matters relating to maintenance obligations (COM(2005)0649 – C6-0079/2006 – 2005/0259(CNS))’, Committee on Civil Liberties, Justice and Home Affairs of the European Parliament (A6-0468/2007) [2007] 41 <<http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//NONSGML+REPORT+A6-2007-0468+0+DOC+PDF+V0//EN>> accessed 1 April 2009.

<sup>30</sup> n 7.

<sup>31</sup> Opinion of the Committee on Legal Affairs on the Proposed Legal Basis (n 29) 41-42.

<sup>32</sup> European Commission (n 19) 5.

<sup>33</sup> Protocol on the position of the United Kingdom and Ireland (UK and Ireland Protocol) [1997] C 340/99.

<sup>34</sup> n 11.

<sup>35</sup> Treaty on European Union (EU Treaty, as amended) [2002] OJ C 325/5.

<sup>36</sup> n 10.

<sup>37</sup> Article 61(c), the legal basis of the Maintenance Regulation, comes under Title IV of the EC Treaty, which covers visas, asylum and immigration as well as civil justice.

<sup>38</sup> UK and Ireland Protocol (n 33), Art 3.

<sup>39</sup> UK and Ireland Protocol (n 33), Art 4.

<sup>40</sup> Baroness Ashton (n 21) 379.

<sup>41</sup> Council preparatory document, ‘Council Regulation on jurisdiction, applicable law, recognition and enforcement of decisions and cooperation in matters relating to maintenance obligations – Adoption’ (17102/08 JUSTCIV 262) [2008], [5].

secure the very existence of Title IV. The rationale for the existence of the Protocol seems quite convincing. The UK is somewhat unique in Europe, geographically and legally speaking, and does not want to lose all control in this politically sensitive area. The opt-in, therefore, gives the UK the required flexibility not to be bound by unpalatable measures, but to participate fully in those that are more acceptable. On this understanding, however, the opt-in is an all-or-nothing option. Either the UK opts-in within three months of the commission's proposal and participates fully in the negotiations and adoption of the act, or it does not. If it does not, then there remains the option to opt in after the instrument's adoption if the UK has a change of heart.

It is this take-it-or-leave-it device that the House of Lords European Union Select Committee recently described as being the UK opt-in:

The effect of this Protocol is that the United Kingdom does not take part in the negotiation and adoption of Title IV measures, and is not bound by them, unless within three months of a proposal for legislation being presented to the Council the United Kingdom notifies the President of the Council that "it wishes to take part in the adoption and application" of the proposed measure.<sup>42</sup>

Initially, the United Kingdom seemed to use its ability not to opt in judiciously.<sup>43</sup> However, such restraint is no longer evident in the field of civil law.<sup>44</sup> In the present writer's submission, the UK Government has ignored the traditional understanding of its opt-in, expounded by the House of Lords Select Committee, preferring instead to use it as a negotiating tactic in order to secure a more desirable outcome.<sup>45</sup>

Such use of the UK's opt-in has been criticised on several grounds, not least by Britain's own MEPs.<sup>46</sup> One of the key concerns is the position this puts British MEPs and negotiators in the Council in. As regards the Maintenance Regulation, it seems that the UK government was always committed to the broad aims of the regulation<sup>47</sup> but had some specific objections, not least about applicable law. Given the UK's ultimate intention to

<sup>42</sup> European Union Select Committee, 'The United Kingdom Opt-in: Problems with Amendment and Codification' HL (2008-09) 55 [2] <<http://www.publications.parliament.uk/pa/ld200809/ldselect/lddeucom/55/55.pdf>> accessed 1 April 2009.

<sup>43</sup> 'In practice, the United Kingdom has opted in to all proposals concerning asylum and civil law and nearly all proposals concerning illegal migration.' S Peers 'Vetoed, Opt-outs and EU Immigration and Asylum Law' 25 October 2004 <<http://www.statewatch.org/news/2004/oct/eu-immig-opt-outs.pdf>> accessed 1 April 2009.

<sup>44</sup> 'The United Kingdom Government has recently made near-systematic use of the 1997 Protocol by not opting in at the outset to three important legislative instruments (on the law applicable to contracts—hereinafter "Rome I"—on matters relating to maintenance obligations and to divorce)...' D Wallis, 'Memorandum to the House of Lords European Union Select Committee' European Union Select Committee, 'The Treaty of Lisbon: an impact assessment' HL (2007-08) 62-II E84 [5] <<http://www.publications.parliament.uk/pa/ld200708/ldselect/lddeucom/62/8010906.htm>> accessed 1 April 2009.

<sup>45</sup> See for example the comments of Baroness Ashton: 'Notwithstanding the formal legal position in respect of the opt-in Protocol we do plan to continue to engage fully and constructively in the development of this proposal in the hope that we contribute positively to the improvement of the measure.' (n 21) 379.

<sup>46</sup> See for example comments of Baroness Ludford before the House of Lords EU Select Committee: "You sort of negotiate and lobby from the outside. You say you are not opting in but you do a lot of lobbying, particularly of MEPs and stuff... it is bizarre" European Union Select Committee, 'The Treaty of Lisbon: an impact assessment' HL (2007-08) 62-II E91; 'Should we actually be working on, speaking on – and, more importantly, voting on – proposals on things that, as they currently stand, will have no application to those who elected us? Others are beginning to question the legitimacy of our position. These opt-outs are democratically untenable and destructive to the coherence of the EU civil justice system.' D Wallis, European Parliament debate on maintenance obligation, (12 December 2007) <<http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+CRE+20071212+ITEM-020+DOC+XML+V0//EN>> accessed 1 April 2009.

<sup>47</sup> 'I should stress at the outset that the UK is very keen to develop and deliver improvements in the way international maintenance cases are handled.' Baroness Ashton (n 21) 379.

participate in the regulation once these concerns had been addressed,<sup>48</sup> perhaps it would have been best to have opted in initially and won the necessary concessions in the negotiations as a fully committed participant, as Diana Wallis suggests:

The UK informed the Council on 27 April that it will not opt in to the proposal, in accordance with a 1997 Protocol. It is incredibly disappointing and there is no clear evidence that the issues that might worry the UK could not be dealt with within the legislative process, as will be the case for all other EU countries.<sup>49</sup>

On this view, participating in the negotiations, whilst not being bound by the outcome, could be seen as an exploitation of the privileged position afforded by the Protocol, and could jeopardise relations with other member states in the Council.<sup>50</sup> In addition, “British MEPs are put in an increasingly untenable situation... they are essentially making legislation for others and not, most peculiarly, for the constituents by whom they are directly elected.”<sup>51</sup>

This view is, however, disputed by both the British Government and its advisers who are involved in the negotiations. As regards council negotiations, the UK government seems to believe it is using the opt-in responsibly in order to simultaneously allow the other member states to make progress unhindered by a British veto, remain engaged in the negotiations, and protect Britain’s interests.<sup>52</sup> As for the position of British MEPs regarding measures the UK has not opted-in to, the Government response seems to be that MEPs represent their constituents not their states, and, given the fact that the UK may opt in in the future, British MEPs should be able to participate fully.<sup>53</sup>

Whilst the UK government’s defence of its use of the opt-in appears quite reasonable, there seems to be a genuine concern that not opting-in initially is becoming the norm. In addition to the practical problems presented by this in negotiations, it is not sending the signal that the UK is fully committed to reform in the area of civil justice. In the present writer’s opinion, both the issue of the UK’s position within the Council and of British MEPs would be resolved if Britain were to have opted in to the Maintenance Regulation at

<sup>48</sup> ‘Our hope would be that when negotiations are concluded that we will be able to accept the proposed Regulation under Article 4 of the Title IV Protocol.’ Baroness Ashton (n 21) 379.

<sup>49</sup> D Wallis, “UK ‘opt-outs’ hit those who are vulnerable” says European Parliament Vice-President Wallis’ 13 December 2007 <[http://dianawallismep.org.uk/news/000545/uk\\_optouts\\_hit\\_those\\_who\\_are\\_vulnerable\\_says\\_european\\_parliament\\_vic\\_eresident\\_wallis.html](http://dianawallismep.org.uk/news/000545/uk_optouts_hit_those_who_are_vulnerable_says_european_parliament_vic_eresident_wallis.html)> accessed 1 April 2009.

<sup>50</sup> ‘The United Kingdom participates actively in the negotiations in Council, but all parties know that it is not bound by the final result. Its European partners are of course in a different position, having no choice but to apply the act once it is adopted. This underlying inequality is arguably not conducive to enhanced trust between Member States, and there is no reason why the United Kingdom could not deal with its concerns through the legislative process, rather than from the sidelines...’ D Wallis (n 46) [9] – [10].

<sup>51</sup> D Wallis (n 46) [12].

<sup>52</sup> See for example Baroness Ashton: ‘Were we to have opted-in, it would have to be on the basis that we would strongly resist features of the proposal which currently seem central within it. Since the proposal is subject to unanimity, that may have required us to block it, on the understanding that under the terms of the Title IV Protocol, the other Member States may, after a reasonable time, proceed without us. That may have been viewed as unhelpful by partner Member States.’ (n 21) 379; see also Lord Mance’s comments regarding Rome I negotiations, which are equally applicable to the Maintenance Regulation: ‘May it not be a bit unfair to say that the United Kingdom failed to engage in those? It did not opt in because it was so engaged and regarded its interests as so engaged, and as far as I can see, it has been very engaged in the negotiations, which have led to a conclusion. One cannot predict ministerial decisions about whether to opt in now but, having negotiated in good faith, a conclusion has been reached which appears to the negotiators to be satisfactory.’ European Union Select Committee (n 46) E91.

<sup>53</sup> See oral evidence given by the Justice Secretary Jack Straw: ‘... although we were not opting in, [British MEPs] could make the instrument more satisfactory because it is bound to have some kind of impact upon us. It would be very, very odd, I think, to say that the Members of the European Parliament, who are elected by voters, should have their ability to vote on individual instruments determined by the position of the government of their host country...’ European Union Select Committee (n 46) E119.

the start, rather than waiting until the negotiations were concluded. However, it is significant to note that this problem may not have arisen at all had the Commission not rushed ahead with its proposal on applicable law against the advice of Member States, which will be discussed more fully in the final section of this paper.

It may seem counterproductive and of little relevance to dwell on an undisputed legal basis and on an opt-in that is enshrined in the Treaty, especially considering negotiations have been concluded and the regulation has been adopted. However, the adoption procedure is of crucial importance to the legitimacy of the instrument and vital lessons can be learned for the adoption of future Private International Law instruments.

Family law is seen as being an especially sensitive area of national law. Therefore, member states are unwilling to cede any sovereignty in this area. Had the Council decided to transfer maintenance obligations to the co-decision procedure with qualified majority voting, member states might have been forced to accept an unwelcome intrusion by the EU into their national family law. This view, however, is based on false assumptions. Firstly, it clings to the notion that family law is somehow uniquely and intrinsically linked to each member state. This assertion is, however, no longer tenable given the increasing internationalisation of family law due to the free movement of persons.<sup>54</sup> In addition, the Council's reticence to transfer maintenance obligations to co-decision shows their unwillingness to recognise the mixed nature of maintenance obligations as pecuniary claims stemming from family law. Therefore, the Maintenance Regulation, which directly affects the lives of EU citizens, represents a missed opportunity to create a more democratically legitimate and inclusive instrument, and this mistake should not be repeated in the future.

As regards the UK's opt-in, the negotiations on the Maintenance Regulation illustrate the potential for abuse of the UK's privileged position. The opt-in is just about acceptable as an escape clause for the UK but it is questionable whether it will continue to be accepted as a strong arm negotiating tactic. Although the UK's position is enshrined in the Treaty, it would be advisable to exercise restraint for the sake of mutual trust and cooperation with the other EU member states.

## Key Features of the Maintenance Regulation

The primary aim of the Maintenance Regulation is to allow a maintenance creditor "to obtain easily, in a Member State, a decision which will be automatically enforceable in another Member State without further formalities."<sup>55</sup> This is achieved by giving effect to the European Council's conclusions reached in Tampere, specifically that which calls for the "abolition of intermediate measures required for the recognition and enforcement in the requested State of a decision given in another Member State, particularly a decision relating to a maintenance claim,"<sup>56</sup> otherwise known as *exequatur*. In doing this, the Maintenance Regulation is intended to replace Brussels I<sup>57</sup> as regards maintenance obligations. In addition to the abolition of *exequatur*, the Maintenance Regulation also

<sup>54</sup> 'A little more uniformity in the law of name, divorce, adoption [etc] can only support free movement and will, at the same time, bring Europe and its citizens closer to one another. Cherishing law as cultural baggage, whatever the circumstances, will lead to intellectual rigidity and isolates us from the benefits which comparative law and harmonisation could bring with them.' W Pintens and K Vanwinckelen, *Casebook European Family Law* (Leuven University Press, Leuven 2001) 15.

<sup>55</sup> Maintenance Regulation (n 1), Recital 9.

<sup>56</sup> Maintenance Regulation (n 1), Recital 4.

<sup>57</sup> n 5.



extends the direct grounds of jurisdiction in Brussels I and provides for a system of administrative cooperation. By describing the key innovations introduced by the Maintenance Regulation and contrasting this with the initial Commission proposal and the European Parliament's proposed amendments, this section will try to assess the strengths and weaknesses of this recent Private International Law instrument.

### **Exequatur**

Perhaps the most striking achievement of the Maintenance Regulation, is the abolition, albeit partial, of *exequatur*. The current regime for the enforcement of decisions relating to maintenance obligations is governed by Article 38(1)<sup>58</sup> of Brussels I, which requires a declaration of enforceability. This requirement is abolished, for the vast majority of member states,<sup>59</sup> by Article 17(2)<sup>60</sup> of the Maintenance Regulation.

Initially, the Commission envisaged that *exequatur* would be abolished between all the member states of the EU.<sup>61</sup> In its proposal, the Commission highlighted the intrinsic link between the abolition of *exequatur* and the harmonisation of applicable law:

... the conflict-of-law rules accompany and facilitate the elimination of 'intermediate measures' at the stage of recognition: the decision is less problematic to accept if it is given in accordance with a law designated according to harmonised rules.<sup>62</sup>

In light of this, the Commission's proposal contained detailed provisions on applicable law, which did not appear in the final Regulation. The reason for this was that by the time negotiations on the Maintenance Regulation had begun in earnest, a Protocol on applicable law had already been concluded at The Hague.<sup>63</sup> Therefore, it was decided to simply adopt those rules wholesale and refer to the Protocol in the regulation.<sup>64</sup>

This was a compromise solution to the divisive issue of harmonisation of applicable law. The compromise in the Regulation reflects the optional nature of the Protocol. Only those member states who will become a party to the Protocol will have harmonised Conflict of Laws rules. In practice, this will be the majority of states. By contrast, those member states that will not be party to the Protocol, notably the United Kingdom, will not have harmonised applicable law rules. Given that the automatic recognition and enforcement of judgments relating to maintenance obligations is only palatable if they are made on the basis of harmonised applicable law rules, the abolition of *exequatur* is only possible between those member states that are contracting parties to The Hague Protocol. Therefore, decisions emanating from member states who are party to The Hague Protocol

<sup>58</sup> 'A judgment given in a Member State and enforceable in that State shall be enforced in another Member State when, on the application of any interested party, it has been declared enforceable there.'

<sup>59</sup> Denmark is not participating in the regulation and *exequatur* will remain for UK decisions.

<sup>60</sup> 'A decision given in a Member State bound by the 2007 Hague Protocol which is enforceable in that State shall be enforceable in another Member State without the need for a declaration of enforceability.'

<sup>61</sup> 'A decision given in a Member State that is enforceable in that Member State shall be recognised and enforceable in another Member State without the need for a declaration of enforceability and without any possibility of opposing its recognition.' Commission of the European Communities, 'Proposal for a Council Regulation on jurisdiction, applicable law, recognition and enforcement of decisions and cooperation in matters relating to maintenance obligations' Com (2005) 649 final article 25 <<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2005:0649:FIN:EN:PDF>> accessed 1 April 2009.

<sup>62</sup> European Commission (n 61) [1.2.2].

<sup>63</sup> Protocol on the law applicable to maintenance obligations (Hague Protocol on Applicable law) (adopted 23 November 2007).

<sup>64</sup> 'The law applicable to maintenance obligations shall be determined in accordance with the Hague Protocol of 23 November 2007 on the law applicable to maintenance obligations (hereinafter referred to as the 2007 Hague Protocol) in the Member States bound by that instrument' Maintenance Regulation (n 1), Art 15.

will not be subject to a declaration of enforceability;<sup>65</sup> whereas, decisions from member states that are not party to the Protocol will be.<sup>66</sup>

In the United Kingdom, this means that a decision given in another member state will be enforceable without the need for a declaration from a UK court. However, decisions given by UK courts will still require a declaration of enforceability from the court in the member state of enforcement. This may not seem fair, but it is the necessary corollary of the UK's rejection of the harmonisation of applicable law rules. Essentially the Brussels I regime, with stricter time limits, will apply to UK decisions.

Although the abolition of *exequatur* was a specific aim within the EU, it is not self-evident that the regime that will apply to decisions emanating from the majority of member states is more advantageous than that which applies to UK decisions. Whilst the abolition of *exequatur* seems desirable in uncontested cases, the Maintenance Regulation allows for greater delays if pernicious litigants want to drag out proceedings because there is the possibility to oppose enforcement in both the member state of origin and enforcement. This was not the case under Brussels I because all challenges against the enforcement of a decision could be brought in one forum. The main improvement is the removal of the possibility of challenging recognition and enforcement on the basis of public policy, the role of which in maintenance claims seems limited.

### ***Jurisdiction***

Another innovation in the Maintenance Regulation compared with Brussels I is the extension of the direct grounds of jurisdiction. Firstly, the Maintenance Regulation replaces the jurisdiction-conferring connecting factor of domicile found in Brussels I<sup>67</sup> with habitual residence.<sup>68</sup> This is typical of recent Private International Law instruments within the EU.<sup>69</sup>

Furthermore, the Maintenance Regulation grants courts some subsidiary jurisdiction, whereas Brussels I did not. Under the Brussels I regime, if the various grounds of jurisdiction had not led to a court of any member state having jurisdiction then the matter was referred to national law.<sup>70</sup> This, however, is not the case for the Maintenance Regulation, which does not refer to national law. This seems desirable in the interests of legal certainty and preventing forum shopping. The Maintenance Regulation instead provides common nationality (or domicile in the case of the UK and Ireland)<sup>71</sup> as a subsidiary ground of jurisdiction.<sup>72</sup> Whilst some may argue that subsidiary jurisdiction is not necessary, this does not seem exorbitant because these connecting factors at least indicate a reasonably sufficient connection with the country. This might be more the case for the UK and Ireland because of the continuing link implied by the concept of domicile but not necessarily by nationality.

In addition to this there is a *forum necessitatis* clause inspired by Swiss law<sup>73</sup> whereby a party who is unable to bring a case in a third state, for example due to civil war, can bring a

<sup>65</sup> Maintenance Regulation (n 1), Art 17.

<sup>66</sup> Maintenance Regulation (n 1), Art 26.

<sup>67</sup> Brussels I Regulation (n 5), Art 2.

<sup>68</sup> Maintenance Regulation (n1), Art 3.

<sup>69</sup> E.g. Brussels II bis (n 24).

<sup>70</sup> Brussels I Regulation (n 5), Art 4.

<sup>71</sup> Maintenance Regulation (n 1), Art 2(3).

<sup>72</sup> Maintenance Regulation (n 1), Art 6.

<sup>73</sup> See Swiss Code of Private International Law, Art 3.

case in a member state with sufficient connection. This is in order to avoid a situation of denial of justice.<sup>74</sup> This is merely an enunciation of a principle that has long been recognised in England. See for example the *Carvalho* case,<sup>75</sup> where the English courts refused to decline jurisdiction and give effect to a choice of court agreement assigning an Angolan court exclusive jurisdiction due to a revolution in Angola which resulted in civil and political unrest.

The commission justifies this subsidiary jurisdiction on the basis that “legal foreseeability is strengthened, and all situations in which a link with the European Community can legitimately be established are covered.”<sup>76</sup> However, as Michael Hellner points out, this is not necessarily the case.<sup>77</sup> There is no mention in the regulation of jurisdiction based on the location of assets for example. This could present a practical problem where a decision is obtained in a non-member state where the defendant has no assets. Unless there is an enforcement agreement between that state and the member state where the defendant’s assets are, the maintenance decision would be ineffectual. This situation could be resolved by conferring jurisdiction on the basis of the location of assets, which would give a court of a member state jurisdiction which otherwise it would not have had. Failing this, a reference to national law may still be necessary because some national laws (e.g. Swedish) provide such jurisdiction.<sup>78</sup>

Another change to the jurisdictional rules of Brussels I is the limiting of the possibility of choice of court agreements. Under Brussels I a choice of court agreement could, in all situations, designate any court. This has been restricted to certain courts and precluded in cases where children under the age of eighteen are concerned. This restriction on party autonomy seems a little unwarranted – the rationale seems to be that parties may choose a court that grants smaller awards of maintenance and thus not be in the best interests of the child. This, however, overlooks the reality that the maintenance would mostly be claimed by one of the parties on behalf of the child. Therefore, they are probably not any more likely deliberately to choose a court that will award lower amounts for child maintenance than they are for spousal maintenance, the latter of which is allowed by the Regulation. Perhaps a sound justification for this difference in treatment is that spouses, being on an equal footing, do not require protection from mistakenly entering into prejudicial choice of court agreements, whereas the interests of the child are so important that they require protection from parents accidentally entering into such agreements. This, however, seems of little relevance, given that there has been no evidence of this kind of theoretical problem ever occurring under Brussels I.

Finally, as regards jurisdiction, it is worth noting the restrictive nature of the *lis pendens* clause of the Maintenance Regulation. Whilst harmonising jurisdictional rules within the European Union is the more important aim of the Regulation, it would have been desirable if it had been more internationalist in its approach. The regulation allows the courts of a member state, seized of a dispute already pending before the court of another member state, to stay proceedings pending the outcome of that case to avoid a clash of judgments.<sup>79</sup>

<sup>74</sup> Maintenance Regulation (n 1), Recital 16.

<sup>75</sup> *Carvalho v Hull, Blyth (Angola) Ltd* [1979] 3 All ER 280.

<sup>76</sup> Commission of the European Communities, ‘Communication from the Commission to the Council and the European Parliament: Commentary on the articles of the proposal for a Council Regulation on jurisdiction, applicable law, recognition and enforcement of decisions and cooperation in matters relating to maintenance obligations’ COM 2006 (206) final 3.

<sup>77</sup> M Hellner (n 27) 348.

<sup>78</sup> Swedish Procedural Code Ch. 10 § 3.

<sup>79</sup> Maintenance Regulation (n 1), Art 12.



However, it remains silent on the issue of cases pending before the courts of non-member states. This could therefore lead to a situation where an applicant could use the courts of the European Union as a means of escaping a negative judgment in a case that is nearly concluded in a non-member state.<sup>80</sup> By not addressing this issue, the provision on *lis pendens* represents a ‘fortress Europe’ mentality, which hopefully will be avoided in any future revision of this instrument and the upcoming revision of Brussels I.

### **Administrative Cooperation**

One of the key innovations of the Maintenance Regulation is the system of administrative and judicial cooperation it creates. This is a novel addition to the Brussels I regime and an improvement on the system of cooperation in the New York Convention. The Maintenance Regulation even seeks to go further in the area of cooperation than the recent Maintenance Convention. Whilst the initial Commission proposal may have been lacking in some respects, the Council has taken on board criticisms of the proposal to create a strong and effective system of administrative and judicial cooperation. Part of the reason for this is that the Regulation follows the Maintenance Convention quite closely and therefore is able to flesh out the initial proposal to a greater extent.

This system of close administrative cooperation is created through the use of central authorities. Although these effectively already existed in most member states under the New York Convention, they will now exist in every member state and have a wider range of functions. One of the main functions of the central authorities will be to facilitate the recovery of international maintenance through the exchange of information. Whilst this was the case under the New York Convention, the Regulation covers a broader range of more sensitive information relating to the debtor. Article 61(2) provides for the transmission of information relating to the location of the debtor and his or her assets, as well as identifying his or her employer and bank accounts. This information, although necessary to allow for effective enforcement, raises several privacy and data protection issues.

The Commission’s proposal did not seem to consider fully the implications of this enhanced exchange of information for data protection. The Council, however, took up the suggestions made by the EDPS in the final regulation. The EDPS expressed concerns that there was not explicit provision relating to the proportionality and necessity of the information that was to be communicated.<sup>81</sup> This is dealt with in Article 61(2), which requires the information to be “adequate, relevant and not excessive.” Whilst both the proposal and the final Regulation list the type of information that can be used, and state that it may only be used in order to facilitate the recovery of maintenance claims, the Regulation goes further by specifying which information can be used at the different stages of recovery.<sup>82</sup> In addition, the final Regulation follows the EDPS’s advice and makes more specific rules as to which authorities within a state are able to provide information to the central authorities.<sup>83</sup>

<sup>80</sup> See for example Case C-281/02 *Andrew Owusu v N.B. Jackson, trading as ‘Villa Holidays Bal-Inn Villas’ and Others* [2005] ECR I-01383, where the ECJ held that, under the Brussels Convention, courts of a member state cannot stay proceedings even if they are already pending in a non-member state.

<sup>81</sup> European Data Protection Supervisor, Opinion of the European Data Protection Supervisor on the Proposal for a Council Regulation on jurisdiction, applicable law, recognition and enforcement of decisions and cooperation in matters relating to maintenance obligations (COM(2005) 649 final) [2006] OJ C 242/20.

<sup>82</sup> Maintenance Regulation (n 1), Art 61(2).

<sup>83</sup> Maintenance Regulation (n 1), Art 61(1).

Part of the concern relating to the use of this information is the extent to which the privacy of the debtor is protected. Under the Commission's proposal, this matter was not sufficiently regulated. For instance, it was not clear the extent to which this information would be available to the creditor. This is dealt with in the Maintenance Regulation through an express obligation to ensure the confidentiality of the information<sup>84</sup> and the prohibition of such information being communicated to the creditor.<sup>85</sup> The Council has also taken on board some other more minor suggestions made by the EDPS: for example, rather than restricting the length of time that information can be kept to a specific period as the Commission suggested, the Regulation instead provides that it must not be kept for longer than is necessary. This addresses the EDPS's concern that a specified period may be too short. Such consultation and amendments are therefore important, and the Regulation that has emerged from this is more respectful of parties' privacy.

Furthermore, many of the rules on administrative cooperation have taken their cue from The Hague Convention, which has resulted in a more practicable and effective instrument. Under the initial Commission proposal, applications had to be made to a court first, which would then be forwarded to the central authorities if the appropriate conditions were met. The involvement of courts in this essentially administrative process would have only served to slow it down. In the final text, the Council opted for a similar solution to The Hague, where applications are made to and processed entirely by administrative central authorities rather than courts.<sup>86</sup> This is a more efficient and speedy system.

A further related development is the inclusion of both the debtor and creditor as potential applicants. Under the Commission proposal, central authorities were only under a duty to assist creditors.<sup>87</sup> This disparity in treatment has been remedied in the final text, which sets out the applications that both creditors and debtors are able to make.<sup>88</sup> An additional criticism of the Commission's proposal is that it does nothing to spur on those central authorities that do not function efficiently. This has, however, been addressed in the Regulation itself,<sup>89</sup> through the adoption of time limits, which are even stricter than those in The Hague Convention.<sup>90</sup>

### ***Enforcement Measures***

In order to ensure effective and rapid access to local justice, the Commission, in its proposal, set out a number of hard hitting enforcement measures that could be instigated from the country of origin rather than the country of enforcement. These were by way of derogation to the principle that enforcement is carried out by the competent authorities in the member state of enforcement. These measures included the garnishment of wages and the freezing of bank accounts. The former allowed the court of the member state of origin to give a direct order for monthly payments in the member state of enforcement. Whilst this measure is available in some member states, this provision would ensure that it is available in all. Furthermore, it would mean that payment was rapid and not subject to the delays that enforcement by the authorities in the member state of enforcement entails. The latter measure was meant only as a temporary step in order to prevent debtors from

<sup>84</sup> Maintenance Regulation (n 1), Art 62(4).

<sup>85</sup> Maintenance Regulation (n 1), Art 62(2).

<sup>86</sup> Maintenance Regulation (n 1), Art 55.

<sup>87</sup> Commission Proposal (n 61), Art 41(1)(b).

<sup>88</sup> Maintenance Regulation (n 1), Art 56.

<sup>89</sup> Maintenance Regulation (n 1), Art 58.

<sup>90</sup> Hague Maintenance Convention (n 2), Art 12.

avoiding the claim by moving their assets. Given how time sensitive this measure is, the ability of the court of origin to issue it is desirable for eliminating delay.

Although these mechanisms are well intentioned, and reflect two of the main means of enforcement available in national law that are suggested by The Hague Maintenance Convention,<sup>91</sup> it is hard to imagine them being implemented in reality. As regards wage garnishment, this involves courts directly in actual enforcement, which might not be an established practice in every member state and therefore the court may not have the necessary expertise. In addition, there is a risk of causing the debtor to pay twice. This is something that the central authorities in the country of enforcement are in the best position to prevent.

As far as freezing of bank accounts is concerned, this provision may go further than is necessary. The EESC suggest that it would be better to allow for the freezing of only part of the money so as to avoid a situation where the debtor does not have money to meet living expenses.<sup>92</sup> The problem is perhaps more fundamental than that because it could be seen as an intolerable intrusion into the sovereignty of the member states. Given that this measure should not be commonly used with regard to maintenance obligations, perhaps it is best dealt with in another context.<sup>93</sup>

Whilst such immediate enforcement measures seem desirable in theory, they are not very feasible in reality. It is probably for this reason that they were not included in the final Maintenance Regulation. Instead, the central authority in the state of origin can request its counterpart in the state of enforcement to take specific measures and the latter will have discretion over which enforcement mechanism to employ. This seems a more acceptable solution because it removes the judicial element, which may not normally be involved in enforcement, and instead places the responsibility on the authorities that are best placed to discharge it effectively.

A further difference in this area between the proposal and the final Regulation relates to the ranking of maintenance claims. In the Commission's proposal, maintenance claims would have ranked above all other claims on the debtor. Whilst this is aimed at ensuring maintenance claims are paid, it does not reflect commercial reality and as such did not appear in the final Regulation.

## Applicable Law

Although the Commission proposed its own set of applicable law rules, the Council decided that the final Regulation<sup>94</sup> should state that The Hague Protocol on Applicable Law<sup>95</sup> should determine the law that applies within the Community. Ironically, although the rest of the regulation builds on and is inspired by work at The Hague, The Hague Protocol seems to take its inspiration from the Commission proposal. This is understandable given that the majority of the contracting states to the earlier conventions on applicable law were EU member states. Many of the non-EU member states at The

<sup>91</sup> Art 34(2)

<sup>92</sup> European Economic and Social Committee, Opinion of the European Economic and Social Committee on the Proposal for a Council Regulation on jurisdiction, applicable law, recognition and enforcement of decisions and cooperation in matters relating to maintenance obligations COM(2005) 649 final — 2005/0259 (CNS) [2006] OJ C 185/35.

<sup>93</sup> E.g. in the context of the Commission's green paper on the attachment of bank accounts. See M Hellner (n 27) 266.

<sup>94</sup> Maintenance Regulation (n 1), Art 15.

<sup>95</sup> n 63.

Hague are less interested in creating harmonised applicable law rules, especially considering that the common law countries do not usually apply foreign law in maintenance disputes.

The general rule in the Protocol, as in the Commission's proposal, is that the law of the habitual residence of the creditor is the applicable law. This is consistent with the previous conventions, and can be justified on the basis that the creditor is usually the weaker party being the one in need of financial support and as such is the subject of the maintenance obligation. Therefore, as the obligation will be discharged in the state of habitual residence of the creditor, it is reasonable that the law of that state should apply.<sup>96</sup>

However, given that the aim is to facilitate the recovery of maintenance claims, if this is impossible under the law of the state of habitual residence of the creditor then, for certain maintenance claims, the law of the forum will apply. The situation is, however, reversed if the creditor chooses to bring the case in the habitual residence of the debtor. In that situation, the law of the forum will apply unless that prevents the satisfaction of the maintenance claim, in which case the law of the habitual residence of the creditor will apply. This is in line with the Commission's proposal, which is justified on the basis that the creditor's choice may be made for considerations of costs or linguistics and should be allowed because there would be no inequality of the parties. Unlike the Commission proposal, which extends this possibility to all maintenance claims, the Protocol only applies these rules to parent-child relationships and other cases involving people under the age of 21.

Just as, under the Maintenance Regulation, the courts of the common nationality of the parties have jurisdiction if no other court has, the law of the state of common nationality applies if the creditor is unable to obtain maintenance under the applicable law rules detailed above. This was also the case under the Commission's proposal. However, it went further and also allowed the law of a country with a close connection to the maintenance obligation to apply. It seems slightly arbitrary to limit this to common nationality because, as the Commission notes, there could be other factors that create a sufficient link. As was the case for jurisdiction, domicile will replace nationality in some countries. It would have been more understandable to specify domicile given the inherent link with that country. Nationality, however, does not necessarily imply any actual connection.

A welcome innovation compared to the previous Conventions that the Protocol brings about is the introduction of party autonomy in the designation of the applicable law.<sup>97</sup> This mitigates the potentially disruptive effect of specifying, as the Protocol does,<sup>98</sup> that upon change of habitual residence, the law of the new habitual residence applies. Although this was not specified in the Commission's proposal, it would have been presumed to be the case. However, neither the Commission's proposal nor the Protocol deals with a situation where the creditor has no habitual residence. This could be problematic where the creditor, upon leaving a country, has lost his or her habitual residence but insufficient time has passed to acquire a new one.

Whilst parties may designate any law applicable for the purpose of particular proceedings, prior agreement as to which law applies to maintenance obligations between

<sup>96</sup> M. Verwilghen, Explanatory Report to the 1973 Convention (15 October 1973) [138] <<http://hech.e-vision.nl/upload/exp123-24.pdf>> accessed 1 April 2009.

<sup>97</sup> Hague Protocol (n 63), Art 8.

<sup>98</sup> Hague Protocol (n 63), Art 3(2).

them is somewhat restricted. This restriction seems aimed at ensuring that a meaningful law is applied such as that of the state of nationality or habitual residence of one of the parties. This is less restrictive than the Commission's proposal, which limited prior designation to the law of the state of common nationality or habitual residence. In addition, the Protocol allows for the law applicable to their property regime or divorce to be designated, whereas the Commission's proposal only allows for the former.

Despite the welcome inclusion of party autonomy, the Protocol, which broadly mirrors the Commission proposal, may be overly restrictive as regards children under eighteen and vulnerable adults. The Protocol in fact prohibits prior designation of the applicable law for such persons. The rationale for this is intuitively understandable as being for their protection so they are not exploited. However, this rests on an assumption that prior agreements designating the applicable law will necessarily disadvantage these parties. In the present writer's submission, this is no more likely the case than with any other party. The increased possibility that these parties' interests might be jeopardised calls for closer scrutiny of these agreements, but not extinguishing party autonomy altogether.

### **Amendments**

The final Regulation is significantly different from the initial proposal. Therefore, it is important to note the changes that have been made in order to understand the influence of the negotiations in The Hague and the European Parliament on the final product. The most significant alteration relates to the provisions on applicable law. Initially the Commission had laid down detailed rules regarding the harmonisation of applicable law. However, these were unacceptable to the United Kingdom, which heavily influenced its initial decision not to opt in. This is quite understandable because common law jurisdictions do not usually apply foreign law in maintenance cases, not least because the costs of proving foreign law are disproportionately large compared to the size of a potential maintenance award. In addition to this, negotiations were underway at The Hague on an additional Protocol on applicable law. Therefore, the Council had to decide whether or not to stick with a second set of applicable law rules, which was not supported by all the member states. In the end, they decided in the Regulation to refer to the Protocol as containing the applicable law rules that would apply in the EU for member states who were contracting parties to it.

The Commission might be criticised for rushing ahead with a proposal containing applicable law against the advice of many member states, when it could have waited until after The Hague negotiations and save some wasted effort. Whilst this is true and should certainly be borne in mind for the future, the work that the Commission did on applicable law allowed the Community to prepare more thoroughly a common position at The Hague than it would have otherwise done.

Another alteration of the Commission proposal is broadening the scope of application of the regulation. Initially the Commission envisaged that the "regulation shall apply to maintenance obligations arising from family relationships or relationships deemed by the law applicable to such relationships as having comparable effects." On the suggestion of



the European Parliament, however, this was extended to cover relationships of affinity.<sup>99</sup> This is seen as one of the benefits of regulating this area within the EU, as closer regional integration allows for a wider scope of application than was possible at The Hague.<sup>100</sup>

An unsuccessful amendment proposed by the European Parliament was intended to restrict party autonomy further as regards choice of court. The European Parliament wanted to go further than prohibiting choice of court agreements for children under eighteen and extend the protection to vulnerable adults too.<sup>101</sup> Given how controversial an issue this was at The Hague, it is unsurprising that agreement was not reached on this point.<sup>102</sup> There does however remain a discrepancy in the position of vulnerable adults as regards party autonomy. Although they are able to conclude choice of court agreements, they are not able to designate the applicable law. This distinction seems rather unwarranted and for consistency's sake should be addressed.

As with its adoption procedure, the substance of the Maintenance Regulation provides lessons to be learned for future Private International Law instruments. Whilst the Regulation makes significant advances on the current regime, there are areas where it could have gone further.

In terms of jurisdiction, it is commendable that the grounds of subsidiary jurisdiction are not exorbitant and do not afford courts within the EU jurisdiction where they manifestly should have none. It is regrettable, however, that this internationally responsible approach was not extended to allow an international *lis pendens* provision. The provisions on administrative cooperation go a long way towards providing a system of effective enforcement but they still do not avoid cases where actual enforcement is not forthcoming, as the Commission proposal did. Whilst this creates an admirable theoretical framework for enforcement, it does not address the reality of the problems that maintenance creditors often face. As regards applicable law, the introduction of party autonomy is a welcome addition. However, the overly restrictive nature of these provisions and the disparity with the provisions relating to choice of court agreements is lamentable.

The most significant achievement of the Maintenance Regulation, and its key aim, was the elimination of *exequatur*. It is regrettable that this has only occurred between most member states and not them all, thus resulting in a two-track Europe for the recognition and enforcement of decisions relating to maintenance obligations. This, however, seemed necessary if all the member states were to participate in the regulation given the inherent link between harmonisation of applicable law rules and the abolition of *exequatur*, and the aversion to the former felt by some.

The link between the substance of the Maintenance Regulation and the importance of its adoption procedure can perhaps be seen most clearly in the amendments that were made to the initial proposal. By taking the opinion of the European Parliament and other institutions that were consulted seriously, the resulting instrument is stronger, more legitimate and more concerned with the rights of the individual parties.

<sup>99</sup> European Parliament Legislative Resolution (n 28), Amendment 16.

<sup>100</sup> P Beaumont (n 3).

<sup>101</sup> 'This Article shall not apply if the creditor is a child below the age of 18 or an adult lacking legal capacity' European Parliament Legislative Resolution (n 28), Amendment 27.

<sup>102</sup> P Beaumont (n 3).

## **Comparison with The Hague**

All the member states of the EU are set to become contracting parties of The Hague Maintenance Convention. Despite this, the EU Maintenance Regulation cannot be said to be superfluous because it achieves things that were not possible at The Hague. The Maintenance Regulation in fact builds on the foundations that were laid down in The Hague Convention. Therefore, it was especially beneficial that negotiations within the EU were put on hold until negotiations in The Hague were concluded.

The aims of the Maintenance Regulation and the Maintenance Convention are broadly similar. Both aim to speed up the collection of international maintenance obligations, and both do this through an enhanced system of administrative cooperation. However, the main area where the Maintenance Regulation was able to go further is the abolition of exequatur, which could not have been achieved at The Hague. This, however, was only possible because of the conclusion of the additional Protocol on applicable law at The Hague.

## ***Family Relationships***

Another area where the Maintenance Regulation was able to go further than the Maintenance Convention was the scope of application. The Convention covers the core cases that form the majority of disputes and represents a minimum consensus. It is only entirely applicable to child support applications for children under the age of 18 and some spousal support applications. There is the possibility for a contracting state of the Convention to apply it to other relationships but that will only be effective if other contracting states make similar declarations. This could lead to disparities between the various contracting states of the Maintenance Convention. In addition, the Convention does not define the term spouse, leaving it open to interpretation by each contracting state. Whilst it is desirable that the Convention did not specify a restrictive definition of the term, this ambiguity could result in a difference in treatment between the different countries. However, perhaps such an outcome was inevitable given the need for unanimity and the desire to secure a high number of ratifications, and the inherent ambiguity should be seen in a positive light.

The Maintenance Regulation, however, was able to level the playing field within the EU and go beyond this lowest common denominator approach. Instead, it covers all family relationships and relationships of affinity. This necessarily covers a wider range of relationships than the Maintenance Convention e.g. same-sex partners and siblings. In addition to achieving this uniformity, the Maintenance Regulation also covers vulnerable adults, which proved to be a divisive issue at The Hague. The fact that closer regional integration resulted in a more progressive and comprehensive coverage of this area than was possible at the international level is one of the main reasons why EU regulation in this field was justified.

Although such harmonisation was necessary to avoid debtors moving to another member state, the internal law of which does not recognise the claim, in order to avoid the claim, the Regulation still required unanimity, and not all the member states recognise the

underlying family relationship. In order to achieve consensus on this and to allay fears that the Regulation would have an unnecessary effect on domestic family law, it included a disconnection clause. This explicitly stated that recognition and enforcement of the maintenance claim did not imply recognition of the underlying family relationship.

Whilst this was necessary in order to achieve consensus it could have been a missed opportunity in terms of European family law. Regulations such as that on maintenance highlight that family law is no longer the sole preserve of domestic law but, given the increase in transnational families, requires a certain internationalisation. Despite this, member states still feel that their sovereignty is violated by harmonisation of their substantive family law. In this regard, incidental harmonisation of Private International Law rules governing the existence of family relationships or even substantive family law may be less intrusive and as such more desirable. However, it is quite possibly the case that the Maintenance Regulation was an inappropriate forum for such harmonisation. As the Commission points out, the harmonisation of applicable law rules renders the harmonisation of substantive law unnecessary.<sup>103</sup>

In addition to laying down foundations, which the EU was able to build on, the negotiations at The Hague also acted as a spur for developments that might not have otherwise occurred within the EU. An example of this is the provisions on legal aid. The Commission's proposal contained less extensive provisions on legal aid than appeared in the final Regulation. This is most likely because provisions that are more extensive were agreed at The Hague, and the EU wanted to match them. Such agreement would not have been reached within the EU itself because of the member states' mixed feelings on legal aid. The impetus for this was provided by the US in The Hague negotiations, which insisted on the inclusion of such wide provisions.

### ***Reverse Subsidiarity***

This analysis highlights the importance of reverse subsidiarity. Measures dealing with international maintenance obligations cannot be adequately dealt with at the national level because of the strong transnational element involved. Therefore, the conditions of subsidiarity are met. However, the EU must consider whether the matter could be more adequately addressed at the global rather than regional level. The above comparison between The Hague Maintenance Convention and the EU Maintenance Regulation suggests that International Family Law and maintenance obligations in particular are best addressed, at least in the first instance, at the more global Hague Conference. After this has occurred, the advantage of closer regional integration means that the EU can build on the strengths of the work at The Hague and go even further.

Despite an initial reluctance, the EU demonstrated, through its actions, that it does take reverse subsidiarity seriously, and proved itself to be a responsible international negotiator. The initial impression from the Commission was that reverse subsidiarity was not a serious concern. This resulted from the way the Commission launched proceedings in the middle of The Hague negotiations, against the advice of the member states, by presenting a proposal, which contained provisions on applicable law. This showed a particular disregard

<sup>103</sup> Commission Proposal (n 60), Art 8.



for reverse subsidiarity because of the likelihood that an additional Protocol on applicable law would be negotiated at The Hague. Inclusion of such provisions suggests an intrinsic preference for Community legislation over international agreement without any basis, especially considering there was a general lack of support for parallel regimes on applicable law.

Despite this, the Commission's proposal did have the fortuitous side effect of promoting a more thoroughly coordinated common position in The Hague negotiations than would have otherwise been the case. If, however, this had been the Commission's intention they should have been more transparent with their reasoning and more explicit in their recognition of the requirements of reverse subsidiarity. In the end, however, this initial 'Fortress Europe' mentality was replaced by a much more internationalist approach, when negotiations in the EU deferred to those in The Hague and did not begin in earnest until after the conclusion of The Hague convention.<sup>104</sup>

It is important, especially in the area of family law, that the EU recognises the value of deferring initially to the more international organisation in The Hague. This has the advantage of providing a foundation upon which the EU can build, as well as the impetus for more extensive reform than would be possible otherwise. The EU seemed to recognise this towards the end. However, in the future, an earlier acceptance of the advantages of reverse subsidiarity would lead to a more efficient use of already over-stretched resources and manpower.

## Conclusion

The Maintenance Regulation is set to replace the existing rules on jurisdiction, applicable law, recognition and enforcement of decisions, and cooperation in matters relating to maintenance obligations in the EU. In all these areas the Regulation represents necessary reform of the current regime. There is, however, always room for improvement: The jurisdictional rules, although considerably extended, still seem somewhat incomplete; the Regulation adopts an overly restrictive approach to party autonomy in relation to both jurisdiction and applicable law; *exequatur* has only been partially abolished resulting in a two-track Europe; and perhaps not enough has been done to give the Regulation real teeth and prevent reticent enforcement by inefficient central authorities.

However, the importance of this latest Regulation on Private International Law within the EU goes beyond the achievements it has made in facilitating the speedy recovery of maintenance claims. The way in which it was adopted has wider significance for the conclusion of future Private International Law instruments: careful consideration should be given to the timing of the Commission's initial proposal; coordination with ongoing international negotiations is crucial; the EU's internal consultation during the legislative procedure can be meaningful and productive, and should be taken seriously; and the United Kingdom must exercise its ability not to opt in with restraint for the sake of its credibility and reputation as a fair negotiator.

Overall, the Maintenance Regulation represents a significant improvement in the recovery of international maintenance obligations and will have a real impact on the

<sup>104</sup> P Beaumont (n 3).

citizens of Europe. More than this though, it is significant because of the internationalist and collaborative approach the EU adopted in relation to parallel developments in The Hague. This is a hopeful sign that the EU is moving away from an introspective approach when it comes to Private International Law and is realising its potential to be a responsible international actor in this field. Such external progress must, however, be mirrored by internal action, for example, by securing an international *lis pendens* clause in any future revision of the Maintenance Regulation and in the revision of Brussels I. Therefore, the Maintenance Regulation, in its substance and adoption, is a qualified success for International Family Law in Europe, and leaves lessons to be learned both from its achievements and from its failings.



# MORAL HAZARD AND HOW IT WAS INVOKED IN THE NORTHERN ROCK CRISIS OF 2007

ANDRÉS CURIA MIRANDA, LL.M.  
KING'S COLLEGE LONDON  
ACURIA@CURIA.CL

**ABSTRACT** In the financial system, the lender-of-last-resort function of central banks and deposit insurance schemes are the foremost source of moral hazard. Both create opposite incentives for regulated individuals who, protected by these regulatory tools, can threaten the very system that they purport to support in the pursuit of profits. In summer 2007, public attention in the UK focused on Northern Rock as its liquidity crisis menaced to impact the entire financial industry. Following the forced support of the bank in September 2007, attention shifted to the allocation of blame with regards to the safety net that public funding provides for both banks and retail depositors in the industry. However, this event was not isolated and only reflects the fundamental consequences that adverse incentives pose on the financial sector.

## Introduction

In the 19th century the accrual of money in banks was the foremost cause which made the money market of England extremely rich and powerful. The concentration of millions in a few lenders and its efficient organisation gave London a significant advantage in the financial market over different trade centres of those times, allowing English banks to grow rapidly.<sup>1</sup>

However, the English banking industry would soon suffer one of the most notorious collapses in its history. This demonstrated that the financial market was, although immense, very fragile. By 1866, Overend, Gurney & Co. was a traditional bank and internationally well known for its credit operations throughout Europe. However, in order to expand its business faster, the bank engaged in hazardous investments, providing loans

<sup>1</sup> Walter Bagehot, *Lombard Street: A Description of the Money Market* (8th edn. Kegan Paul, Trench & Co. London 1883).

against poor collateral which soon led to its insolvency.

Consequently, people were frightened of losing their savings and depositors became a rout. A run soon started upon every bank in Lombard Street. After desperate calls for assistance, the Bank of England agreed to lend millions to save the banks and the run came quickly to a halt.<sup>2</sup>

Although numerous lessons were learnt about the intricacies of the credit system, in September 2007 the United Kingdom suffered a new bank run, with hundreds of depositors trying to withdraw their money from the Northern Rock bank. Despite taxpayers' complaints, the Bank of England had to announce its full support in order to stop the run and the threat of contagion to other financial institutions.<sup>3</sup>

In both cases (only two bank runs severely affected the UK in more than a century) the central bank had to act as a "lender of last resort" in order to provide emergency liquidity support and restore confidence in the financial system. However, the existence of this scheme reduces the cost of risk taking. Subsequently, banks will take on more risk in their investments under this safety net arrangement, increasing the likelihood of a liquidity crisis.<sup>4</sup> Similarly, users of financial services may be induced to take more risk in the pursuit of higher profit, since deposit insurance schemes cover their losses in case of a bank failure, which at the same time may cause banks to invest more recklessly.

Even though the lender of last resort facility and the deposit insurance scheme constitute two of the more significant regulatory tools to prevent financial crisis, they simultaneously place adverse incentives that may threaten the very system that they seek to protect. This phenomenon is called 'moral hazard' and stems from the insurance industry's analysis, with which both facilities share numerous characteristics.

This research essay addresses the phenomenon of moral hazard, particularly in respect of lender of last resort schemes and deposit insurance in the financial system. The first section of this paper starts with a brief description of moral hazard in the insurance industry, followed by a portrayal of the main contributions of Thornton and Bagehot, who are often credited with establishing modern lender-of-last-resort and moral hazard theories in banking.

In the second section, this paper analyzes the different adverse incentives that emergency liquidity support and depositor protection can place onto the banking system. Furthermore, it discusses three diverse approaches that have been suggested to tackle the problem, i.e. constructive ambiguity, the 'too big to fail' assumption and the alignment of decision makers' incentives with regulatory objectives. These aspects of the essay will assist in understanding how moral hazard was invoked in the Northern Rock crisis of 2007, which is dealt with in the last section, together with an outline of the Special Resolution Regime (SRR) recently introduced by the Banking Act of 2009.

<sup>2</sup> Geoffrey Elliot, *The Mystery of Overend & Gurney: Adventures in the Victorian Financial Underworld* (Methuen Publishing, London 2006).

<sup>3</sup> George Walker, 'Credit Crisis: Regulatory and Financial System Reform' (2007) 10(10) JIBFL.

## The Moral Hazard Concept

### *Description of Moral Hazard in the Insurance Industry*

The moral hazard problem has been considered a part of the economics of information, particularly in contracts and authority relationship in economic organisations, where one party's lack of information about the other party's behaviour tends to pose adverse incentives that, subsequently, are detrimental for both of them.<sup>5</sup>

Even though moral hazard pervades in many contractual arrangements, the phenomenon has its roots in the insurance industry. This industry became aware that, when a party is fully insured, the costs of care are high and the insured person cannot be accurately monitored by the insurer, his behaviour may, consequently, change after the policy has been purchased, to the extent that he can affect the likelihood or magnitude of the event that triggers payment.

Therefore, it is submitted that owning an insurance policy increases the risk that policyholders will suffer losses, because it tends to lessen the incentives for the insured party to take actions in order to prevent them, especially when the insurer is unable to supervise its conduct.<sup>6</sup>

### *The Modern Theories of Moral Hazard In The Lender Of Last Resort Facility: Thornton And Bagehot*

In the lender of last resort function of central banks, whereby the authority must provide liquidity to financial institutions when they cannot be demanded in an alternative source, one finds numerous characteristics of the phenomenon. Thornton and Bagehot developed the key elements of the modern doctrine of the lender of last resort in England which is still emphasised until today.

In depicting its distinguishing characteristics, Henry Thornton was the first who described the moral hazard that constantly arises from the lender of last resort.<sup>7</sup> After the deep banking crisis of 1793, which made many country banks fail, Thornton, in his seminal work *An Enquiry into the Nature and Effects of Paper Credit of Great Britain*, portrays the different micro and macroeconomic aims that public banks must face when dealing with liquidity provisions. In this vein, Thornton asserted that the lender of last resort is a monetary rather than a banking function that seeks to maintain the entire quantity of money and the level of economic activity in times of credit distress.<sup>8</sup>

Even though Thornton recognised the danger that a risk of one particular bank can pose on the entire industry, he explained that it is not the role of central banks to stop the initial failure, but to inject liquidity into the market and minimize its secondary repercussions. If any one bank fails, Thornton added, "A general run upon the neighbouring ones can take place, which, if not checked in the beginning by pouring into the circulation a large quantity of gold, leads to very extensive mischief."<sup>9</sup> Therefore, the lender of last resort's

<sup>4</sup> Robert Sollow, 'On the Lender of Last Resort', in Goodhart & Illing (eds.) *Financial Crisis, Contagion, and the Lender of Last Resort* (Oxford University Press, Oxford 2002).

<sup>5</sup> Jackson, et al. *Analytical Methods for Lawyers* (Foundation Press, New York 2003).

<sup>6</sup> Robert Pindyck and Daniel Rubinfeld, *Microeconomics* (3rd edn. Prentice Hall, New Jersey 1995).

<sup>7</sup> Thomas Humphrey and Robert Keleher, 'The Lender of Last Resort: A Historical Perspective' (Goodhart and Illing (Eds) (n 4)).

<sup>8</sup> Henry Thornton, 'An Enquiry into the Nature and Effects of the Paper Credit of Great Britain' (Goodhart and Illing (Eds) (n 4)).

<sup>9</sup> Thornton (n 8).

primary responsibility is to the general interest and not to a specific bank, which may require a special liquidity provision because of the excessive risks its managers have taken.

Furthermore, Thornton asserted that the facility should not imply that it would become the Bank of England's task to relieve every distress provoked by the recklessness of banks, for it "might encourage their improvidence" or misconduct.<sup>10</sup> In other words, the very provision of lender of last resort may create incentives that encourage laxity in the lending practice of individual banks, which in turn would persuade other banks to take speculative risks without reservations of the consequences. Hence, the individual improvidence of banks must be punished by letting them suffer the natural consequences of their fault, and only if the financial effects of such punishment threaten to become widespread, should the lender of last resort step in vigorously.<sup>11</sup>

Thornton also argued the belief that economic welfare is unavoidably affected whenever a bank fails. The author asserted that this argument would provide every large bank, regardless of how it is managed, with an instant justification for support. He considered that often the public interest might be better served by the demise of inefficient institutions, for the resulting improvements in resource allocation may outweigh any adverse side effects of the failure.<sup>12</sup>

Likewise, in *Lombard Street: A Description to the Money Market*, Walter Bagehot emphasised that the lender of last resort should not be a permanent practice of central banks but rather a temporary emergency measure applicable only when the whole economy is at stake. Hence, it should not aim at preventing key banks from failing as a consequence of its poor management, but to avoid a resultant wave of failures disseminating through the system. If banks are bad, Bagehot asserted, "they will probably become worse if the Government sustains and encourages them. The cardinal maxim is that any aid to a present bad bank is the surest mode of preventing the establishment of a future good bank."<sup>13</sup>

Nevertheless, the author believed that in case of panic caused by a bank in considerable distress, funds might be provided by the central bank but at a high interest rate; so high that the rate would function as a 'heavy fine' for borrowers.<sup>14</sup> This penalty rate would provide an incentive for banks to exhaust all market sources of liquidity and even develop new techniques of money management before coming to the lender of last resort. According to Bagehot's analysis, this penalty rate would also furnish the financial system with a test of soundness of distressed borrowers. If the penalty rate on alternative resources of funds is set above the market rate, this would encourage illiquid banks to turn to the market first. Success in obtaining liquidity at the market rate would indicate that lenders appraise these borrowers to be sound risks. In contrast, resort to the central bank would suggest weaknesses in the borrowing institutions, indicating that the bank may be unable to borrow in the market at the lower rate. Fearing default, lenders may demand a risk premium in excess of the difference between the market rate and the penalty rate. This would force the bank to close, to arrange a merger with other institutions or to resort to the central bank's lending facility.<sup>15</sup> However, the central bank support was not meant to substitute prudent bank practices. Hence, even at a penalty rate the last resort lending should not be given to any insolvent firm. In this vein, Bagehot suggested that the majority

<sup>11</sup> *Ibid.*

<sup>12</sup> Humphrey and Keleher (n 7).

<sup>13</sup> *Ibid.*

<sup>14</sup> Bagehot (n 1).

<sup>15</sup> *Ibid.*

<sup>16</sup> Humphrey and Keleher (n 7).



to be protected are the ‘sound’ people, those who have good security to offer. Therefore the lending should be made on liquid type of collateral or on ‘all good banking securities’ that are indisputably good in normal times.<sup>16</sup>

In light of the Overend, Gurney & Co crisis of 1866, Bagehot pointed out that both these rules for last resort lending, i.e. the penalty rate and the need for borrower’s good collateral, must be set out clearly from the very beginning of a crisis. Hence, showing that the central bank is able and willing to stop a panic, but at the same time pointing out that when the exigency had passed “it might let the offending banks suffer.”<sup>17</sup>

## **Adverse Incentives of Safety Net Arrangements**

### ***Emergency Liquidity Provision and Moral Hazard***

The principles of the lender of last resort function of central banks put forward by Thornton and Bagehot still hold today, despite the constant development of the industry. Nevertheless, the moral hazard phenomenon continues to pervade within the financial system because of the information asymmetry and the difficulties to eradicate the incentives that safety net arrangements adversely place on financial institutions.

In a normal interbank market, the surplus liquidity in some banks is transferred to temporarily illiquid institutions for the relevant interest rate, which reflects the risk of the lending. Therefore, the incapacity of a bank to borrow funds through the market, as Bagehot suggested, might indicate that it is insolvent or failing. However, because the market has only access to incomplete information regarding its participants, under certain circumstances solvent banks may be unable to borrow and the distinction between illiquid and insolvent institutions cannot be drawn in the short time.<sup>18</sup>

Central banks’ liquidity provisions will then attract illiquid institutions, either for temporary capital support through the discount window – whereby it provides credit lines secured by good collateral – or for risk capital for a longer period. Yet, gathering information necessary to discriminate between types of banks is complex and costly, even for central banks, which in turn might not have time to verify whether or not a bank is solvent. This is especially so when banks ask for emergency capital injections. Besides, if the central bank provides support to a bank that is revealed later to be insolvent, it will incur not only a direct financial cost, but also a reputational cost for giving support to an imprudent institution. All in all, central banks must weigh the benefits of preventing contagion and panic now against the cost of inducing riskier activity in the banking system later.<sup>19</sup>

Potential adverse incentive structures can thence be identified in these expectations of bailouts or support from central banks, displaying a moral hazard problem on the part of the financial institutions, particularly on its owners and managers.

Even at a penalty rate, financial support from the central bank may induce bank managers and shareholders to pursue higher risk-reward strategies so as to make the best use of the subsidy implicit in such a rescue, and even to ‘gamble for resurrection’ in time of

<sup>16</sup> Bagehot (n 1).

<sup>17</sup> *Ibid.*

<sup>18</sup> Xavier Freixas, et al. ‘Lender of Last Resort: A Review of the Literature’ (in Goodhart and Illing (Eds) (n 4)).

<sup>19</sup> Ross Cranston, *Principles of Banking Law* (2nd edn. Oxford University Press. Oxford 2002).



crisis. Thus, the higher the penalty rate set forth, the riskier the behaviour of the managers and shareholders of the insolvent institution might be, increasing the likelihood of potential losses to a central bank simultaneously. In this vein, the very provision of an interest rate for financial support much higher than market levels, although it might lessen the moral hazard problem as Bagehot suggested, can intensify the bank's crisis and send a signal to the market that precipitates a bank run.

### ***Constructive Ambiguity***

The moral hazard problem implicit in financial supports and bailouts may be restricted by placing an element of uncertainty about the probability of public support before a financial crisis begins. This 'constructive ambiguity' would therefore compel a bank's managers and shareholders to invest cautiously, since they would not be able to judge a priori whether the bank will be assisted or not. Consequently, under this permanent uncertainty, banks would ignore the exact timing of the potential assistance and the conditions and penalties attached to any particular last resort intervention. Thus, it reduces the incentives for excessive risk taking.<sup>20</sup>

Nevertheless, while helping to keep a tight rein on moral hazard, the constructive ambiguity may threaten the stability of the financial system that it seeks to protect. Distinguishing the exact point of departure of a financial turmoil is a risky evaluation in itself. Moreover, a financial crisis can both be triggered and deepened by the lack or the delay of liquidity support, the effects of which can rapidly spread to the rest of the economy.

As the recent crisis showed, the delay of safety net arrangements, inadequate information disclosure and the high discount window's penalty rate charged by the Bank of England intensified the shortage of resources, especially after the international interbank market froze in 2007.<sup>21</sup> The Northern Rock crisis also indicates, as analysed further in this paper, that a delay in satisfying the immediate liquidity demand from a central bank can subsequently impede the timely restructuring of a firm.<sup>22</sup>

### ***The 'Too Big to Fail' Assumption***

If, as Thornton claimed, this economic menace rationale of bank failure provides financial institutions with an instant excuse for liquidity support, regardless of how recklessly it has been managed, which bank should be rescued in order to ensure financial stability? One frequent answer is to rescue only large financial institutions. Particularly under a financial turmoil, regulatory authorities have strong incentives to prevent the bankruptcy of large, highly interconnected financial firms, because of the risks such a failure would pose to the financial system and to the broader economy.

Furthermore, these incentives include preventing the threat of contagion to the international financial market, due to securitisations, collateralised loan obligations and other instruments through which large banks regularly trade credit risk, market risk and interest rate risk.<sup>23</sup> This means that its failure may create a liquidity drain and domino effect within the international financial market and consequently affect the economy as a

<sup>20</sup> Freixas (n 18).

<sup>21</sup> Walker (n 3).

<sup>22</sup> House of Common Treasury Committee, 'The Run on the Rock. Fifth Report of Session 2007-2008' HC (2008) Volume I.

<sup>23</sup> Walker (n 3).

whole. A financial institution thence is deemed to be ‘too big to fail,’ which would lead to an immediate support in order to avoid any contagion and to prevent the distress of thousands of depositors, as a matter of public policy rather than of market regulation.<sup>24</sup>

Nevertheless, the belief that any particular financial institution cannot be allowed to fail because of its dimension and influence, both in the market and in the payment system, has many adverse effects and perverse incentives in itself. In this regard, Charles Bernanke suggests that this too big to fail assumption persuades excessive risk-taking by firms, hence reducing market discipline by the ‘imitation effect’ amongst companies. Besides, according to Bernanke it provides an artificial incentive for firms to grow, in order to be considered as too big to fail and obtain the liquidity support or bailouts in times of crisis, at taxpayers’ expense. Further, it creates an unlevelled playing field with smaller firms, which may not be regarded as having implicit financial support.<sup>25</sup> In this vein, Mervyn King added that a large institution does not require a different approach from the regulatory body, but rather only more effort from the authority to deal with its crisis. Thence, although it is much harder to resolve problems with big banks than with smaller ones, it is not a question of principle, but of scale.<sup>26</sup>

However, the provision of unprecedented publicly financed support to halt the current crisis, especially to systemically significant firms, has crystallised the concern that many banks are indeed too big to fail. In this regards, the Financial Services Authority (FSA) has recently suggested that the largest, systemically important banks may be subjected to higher capital and liquidity requirements, along with restricting the range of activities that these financial institutions can engage in.<sup>27</sup>

There is a fine balance in any decision to rescue a bank, in which dealing with the moral hazard problem plays an important role for preventing future crises. But unless some banks are permitted to fail, the adverse incentive of government’s safety net will still pervade the financial system.

### ***Aligning Incentives and Objectives of Decision Makers***

Alternatively, one form of addressing the moral hazard problem when designing a regulatory framework, and before a financial crisis takes place, is to adjust the incentives of the different stakeholders in the decision making process with the objectives set by regulators. Hence, a central role of policy makers is to induce managers and shareholders to perform in a way consistent with the aims that are set for financial regulation. These aims are systemic stability and depositor protection. This is certainly so because these may not always be in the direct interest of either managers or owners of banks.<sup>28</sup> Shareholders of banks are entitled to monitor managers’ behaviour and thereby may reduce the opportunities of excessive risk taken in its business. However, current capital adequacy requirements allow banks to operate with low capital base, which creates moral hazard since the less amount owners have to lose, the more prone they are to ignore excessive risk taking in the pursuit of profit.<sup>29</sup> If shareholders absorb the risk of a bank, more regulatory

<sup>24</sup> Jan Dalhuisen, *Dalhuisen on Transnational and Comparative Commercial, Financial and Trade Law* (3rd edn. Hart Publishing, Oxford 2007).

<sup>25</sup> Ben Bernanke, ‘Financial Reform to Address Systemic Risk’ (Speech at the Council on Foreign Relations 2009) <<http://www.federalreserve.gov/newsevents/speech/bernanke20090310a.htm>> accessed 15 July 2009.

<sup>26</sup> House of Common Treasury Committee (n 22).

<sup>27</sup> FSA, *A Regulatory Response to the Global Banking Crisis* (Report). (London 2009).

<sup>28</sup> David Llewellyn, *Some Lessons for Bank Regulation from Recent Crises* (DNB Staff report 2000, N° 51. De Nederlandsche Bank NV. Amsterdam 2000).

capital may be required in order to induce them to take precautions and to supervise the performance of managers. In addition, it may be requested to banks' owners to inject risk resources to the business so as to structure an internal insurance fund. They can lose these funds in the event that the firm makes hazardous investments, thus reducing the moral hazard. In doing so, regulatory agencies will also protect the financial aim of systemic stability, since liquidity crises within the market may be ameliorated by banks' internal sources of funds.

On the other hand, if most of the business' decision making is the responsibility of the bank manager, then its incentives must be also aligned with the financial system's ends. This also requires a high degree of professionalism in bank managers as well as consciousness of the different risks that investments may place both on the firm and on the financial system.

In this regard, Dobson emphasised that in contractual environments with moral hazard, reputation can successfully perform its function as an implicit contractual enforcement mechanism for agents. Dobson illustrates this through a rather similar repeated prisoners' dilemma game. Under a constant contractual relationship the reputation building behaviour tends to make an agent choose to sacrifice short run income in the expectation of greater long run revenue. Therefore, Dobson concluded that the agent's desire to build and maintain a reputation places an incentive to render a high quality service even if it yields less profit, which tends to mitigate the moral hazard problem.<sup>30</sup>

Nonetheless, Dobson's results assume a perfect information symmetry, whereby all parties involved (in his principal and agent game) are aware of the relevant data and make their decision accordingly. Yet, asymmetric information is one of the market failures that also affect the financial industry, having regulatory authorities to insist upon effective systems of internal oversight within financial institutions.<sup>31</sup> However, this should not mean that financial regulation may be strengthened, making these systems compulsory amongst regulated institutions. Rather, it suggests that authorities may agree with regulated firms in a protocol that enables them to better audit managers' behaviour periodically, since shareholders and decision makers within firms are in a better position than financial authorities to analyse the banks' performance.

In addition to establishing an internal audit process, it is equally important that managers' remuneration packages relate not only to the firm's performance, but also to regulatory compliance. Within the banking industry, bonuses for managers are typically based on the performance of the company, which traditionally has led to moral hazard problems, since it encourages excessive risk taking in the pursuit of high bonuses.

To illustrate, in the years preceding the crisis of 1930s, Albert Wiggins, president and chairman of the Chase in New York, who also had a reputation as a rather scholarly man, received \$275,000 in compensation as head of the bank.

These bonuses induced Wiggins to engage in speculative operations which, in turn, made the bank suffer severely in the aftermath of the depression.<sup>32</sup> Likewise, Alistair Darling has recently emphasised that many of the problems of the recent crisis were caused by the fact

<sup>29</sup> Dalhuisen (n 24).

<sup>30</sup> John Dobson, 'Moral Hazard, Adverse Selection and Reputation: A synthesis' (1993) 19(6) *Managerial Finance*.

<sup>31</sup> Llewellyn (n 28).

<sup>32</sup> John K Galbraith, *The Great Crash 1929* (Penguin Books, London 1954).

that some bankers were incentivised to take risks that they did not fully understand. Consequently, the manager's bonus culture has had disastrous effects for the economy as a whole; all the more so in nationalised banks such as the Royal Bank of Scotland, where managers received important bonuses despite its failure.<sup>33</sup>

However, in curtailing the effects of manager's pay plans to reduce the moral hazard, regulatory agencies and governments face a new dilemma. On the one hand, governments and regulatory agencies aim to protect and maintain the financial stability, reducing the incentives of short-termism and reckless behaviour within the financial industry. On the other hand, being the largest shareholder in many banks, they require the highest qualified managers, who, if not compensated accordingly, may have incentives to move to other financial centres abroad or to different industries where careers are free from regulation.

As a result, reducing the moral hazard phenomenon is cumbersome and it typically cannot be eliminated in the financial market unless shareholders and managers lose if the firm fails. Yet, internal audit committees within financial institutions may contribute to ameliorate its effects by decreasing the information imbalance and enhancing supervision over manager's behaviour. In one sense, the less the principal monitors, the more latitude is given to the agent in choosing his action and hence the weaker the element is of authority.<sup>34</sup>

In addition, as the FSA introduced the Code of Best Practice in 2009, regulatory bodies may agree with financial firms and international regulators on protocols that review remuneration policies to decrease the incentives for excessive risk taking in the financial system.<sup>35</sup>

### ***Deposit Insurance and Moral Hazard***

After the Great Depression of the 1930s, the US government introduced a deposit insurance programme that sought to protect consumers from bank failures and to prevent any systemic risk caused by a large withdrawal of funds.<sup>36</sup> Similarly, in the UK, the Financial Services and Market Act (FSMA) 2000 set out the Financial Services Compensation Scheme (FSCS), a programme that replaced eight previous compensation arrangements, with the purpose of assisting depositors when they are unable to satisfy their claims against financial institutions in distress, thereby lessen the risk of a bank run.<sup>37</sup>

Nevertheless, it has been claimed that deposit insurance programmes create moral hazard on both depositors and financial institutions. Given that depositors will be protected against any failure, these schemes counteract the consumers' incentives to oversee excessive risk-taking banks. Furthermore, under some circumstances depositors may even be encouraged to look for risky banks on the grounds that they can provide a higher profit in return for their secured investments.<sup>38</sup>

Similarly, because risk is subsidised, insured banks may be induced to take more risks in their business since they are not obliged to pay the full risk premium on insured deposits. Deposit insurance schemes, therefore, enable banks to make riskier loans on a large scale

<sup>33</sup> I Oakeshott, 'News Review Interview with Alistair Darling' Sunday Times (London 16 August 2009).

<sup>34</sup> D Mookherjee, 'The Theory of Economic Organization: Aspect of Adverse Selection and Moral Hazard' (DPh thesis, LSE, 1982).

<sup>35</sup> See FSA, *Annual Report 08/09*, <[http://www.fsa.gov.uk/pubs/annual/ar08\\_09/ar08\\_09.pdf](http://www.fsa.gov.uk/pubs/annual/ar08_09/ar08_09.pdf)>

<sup>36</sup> Galbraith (n 32).

<sup>37</sup> FSMA Act 2000 Part XV. Para 213.

<sup>38</sup> Cranston (n 19).

than they would otherwise, and transfer the cost of deposit protection to others.<sup>39</sup> Consequently, this incentive might be particularly harmful to the financial system, since only the more speculative banks would be able to attract large sums of money, forcing more conservative financial institutions to act accordingly in order to compete for a market share.

However, the claim that depositor protection poses adverse incentives on retail investors rests on rather optimistic assumptions that consumers make their decisions following a reasoned assessment of financial products and institutions, and that they are capable of accurately evaluating the information available in the market.<sup>40</sup> Even if this information is made available by regulation, it will generally require some degree of expertise to understand it, knowledge that needless to say not all depositors possess. Given this manifest asymmetry of information and the less costly option of withdrawal of deposits, in practice only a minimum function of monitoring can be allocated to retail depositors, who have scarce opportunities to be vigilant on the sophisticated financial business.

Therefore, an important role of regulatory agencies is to oversee the performance of financial firms on behalf of customers and protect them in case of any failure. This is particularly so in the case of retail depositors who entrust the banking system with a great proportion of their savings.<sup>41</sup>

As regard the FSCS, the compensation introduced a co-insurance as an attempt to share the risk of bank failure with depositors. According to the FSMA 2000, the programme limited the amount of insurance coverage under the general principle that consumers should take responsibility for their decisions.<sup>42</sup> Consequently, the moral hazard problem would be diminished, for the depositors would have the incentive to take precaution on which financial institution they invest, particularly when these investments exceed the amount insured by the scheme.

Nevertheless, the approach adopted by the FSCS in this protection consisted in giving depositors priority over others creditors on a bank's insolvency. This means that consumers might have to wait months before obtaining payment of the deposits guaranteed.<sup>43</sup> This in itself counteracts the effects that the scheme seeks to achieve, since it creates a strong motivation for depositors to join a bank run and withdraw their money, despite the insurance.

Subsequently, two conclusions can be drawn. First, as regards the FSCS, moral hazard disappeared. It will not be irrelevant for a depositor where to place its savings, because in case a financial institution defaults, the deposit will not be repaid promptly, regardless of the amount guaranteed by the programme. Depositors will thereby need to be vigilant of the financial market in order to keep their savings, depriving the financial system of the stability that regulation seeks to attain. Second, authorities and regulatory bodies need to be liable of constantly informing, advising and warning consumers about the performance of regulated financial firms in a manner sufficiently simple for the understanding of the majority of consumers of financial services.

As a result, these information duties of financial behaviour, together with thorough

<sup>39</sup> Llewellyn (n 28).

<sup>40</sup> Bronwen Morgan and Karen Young, *An Introduction to Law and Regulation* (Cambridge University Press, Cambridge 2007).

<sup>41</sup> Llewellyn (n 28).

<sup>42</sup> FSMA 2000 s.5(d). Prior to 1st October 2007, the FSCS covered 100% of the first £2,000 of deposits, but only 90% of the next £33,000.

<sup>43</sup> Cranston (n 19).



supervision by regulatory agencies, would not only contribute to educate financial services' consumers, but would also lessen moral hazard, since banks would be prone to demonstrate the soundness of their business in order to persuade consumers and attract a larger share of the financial market.

### **Moral Hazard and the Northern Rock Crisis of 2007**

Formerly a building society, the Northern Rock plc was a large bank in the UK, whose core business was the provision of residential mortgages, accounting for 7% of all mortgage lending in the country. Its secured investment, strong balance sheet and constant growth led to Northern Rock entering the FTSE 100 in September 2001.<sup>44</sup>

However, six years later, in September 2007 Northern Rock suffered one of the deepest crises that threatened the financial market, as a bank run started over its few branches. In order to halt the fear and prevent contagion, Parliament had to pass emergency laws to rescue the bank and to take it to public ownership.<sup>45</sup> Still, this event also proved to be a crisis for the financial system, since it revealed the weaknesses of safety net structures and regulatory agencies' supervision.

Moral hazard was part of the analysis of this crisis, due to the adverse incentives that safety net facilities pose onto the financial system's participants. In the following, this essay examines how moral hazard was invoked in the crisis of the Northern Rock of 2007 in light of the House of Common Treasury Committee's enquiry on the issue, particularly in respect of the aspects already analysed in this paper, i.e. the safety net created by the lender of last resort facilities and deposit insurance schemes. This section of the essay ends with an outline of the changes that the Banking Act of 2009 introduced in regards of these features.

### **The Beginning of Northern Rock's Fall**

In order to expand its business, Northern Rock's managers modified the structure of its liabilities, borrowing more funds from the wholesale markets through securitisations rather than funding through retail deposits. Moreover, as much as 75% of the bank's funding came from short-term borrowing, mainly through the issuance of asset-backed securities rather than long-term deposits, particularly in the first half of 2007.<sup>46</sup> Eventually, the interbank interest rate rose for Northern Rock's lending and its share price fell more than any bank at that moment. Hence, it showed that the market was already giving signals about the risks that the bank had taken.<sup>47</sup>

Furthermore, the American sub-prime mortgage market caused a global shock to the financial system, which led to the bank's funding resources to freeze simultaneously in August 2007. Northern Rock's liquidity crisis became evident, and the possibility of the Bank of England giving emergency support started to be considered by the board of the company.

Once the credit crunch began, different banks approached the Bank of England asking for additional liquidity at no penalty rate and against weaker collateral, as other central banks were providing this with the objective of reducing the effects of the credit crunch.

<sup>44</sup> G Davies, 'The Roots of Moral Hazard' *The Guardian* (London 15 September 2007).

<sup>45</sup> Editorial, 'Fail-safe: The problems of winding up huge financial firms' *The Economist* (London 20 August 2009).

<sup>46</sup> W Buiter, 'Immoral Hazard and Northern Rock' *The Financial Times* (London 19 February 2008).

<sup>47</sup> House of Commons Treasury Committee (n 22).

However, the Bank of England did not agree with the suggestions for additional measures, on the grounds that providing this support would place a risk of moral hazard within the financial system.<sup>48</sup> In this regard, the Governor of the Bank of England pointed out that should the bank act and effectively provide extra liquidity against weaker collateral, markets would take it as a signal that the central bank would always rescue them should they take excessive risk and get into difficulties. This is particularly so since the liquidity was provided at little or no penalty rate. Such a signal would lead to even more risk taking and the next crisis would consequently be greater than it would otherwise have been. Yet, Marvyn King, Governor of the Bank of England, emphasised that if the problems that the banking system were facing had been the result of a completely different cause, such as a terrorist attack, the Bank would inject liquidity at absolutely zero cost, for that would not be the result of the risk that banks themselves took.<sup>49</sup>

On the other hand, the lack or delay in liquidity support can deepen the crisis and also expose the system to risk that may reach the economy as a whole. Callum McCarthy, Chairman of the FSA referred to this situation as the problem of “damaged innocent bystanders.” When liquidity is drying up it may affect not only people who have played a part in irresponsible behaviour, but much more widely in terms of other people who can possibly be harmed by that event.<sup>50</sup>

Consequently, these two different approaches demonstrate that there is a fine balance between limiting moral hazard on the one hand and to satisfy the banking system’s demand for additional liquidity on the other. This is certainly true when central banks and regulatory bodies need to deal with individual firms in distress.

When Northern Rock requested extra liquidity in August 2007, the Bank of England attached more weight to the moral hazard argument than to risks of bank failure and contagion. In this regards, the FSA and the British Bankers’ Association agreed that if extra liquidity had been made available to Northern Rock earlier, it would not have subsequently needed to apply to the lender of last resort facility.

### ***The Run on the Rock***

By the third of September 2007, Northern Rock abandoned its attempts at securitisations and the pursuit of a ‘safe haven’ of a takeover by a major bank and formally discussed a “backstop” facility with the authorities. One week later, on the tenth of September, the arrangement was agreed among the tripartite authorities, i.e. the Bank of England, FSA and the Chancellor of the Exchequer, when the necessity of lender of last resort facility to the bank was already more than evident.

This committee also agreed with Northern Rock that the emergency liquidity support should be announced on September 17, thence should wait another week to go public. However, the ‘covert operation’ did not work as planned. The emergency assistance to the bank soon became known to many market participants and the outlines of the operation were reported by BBC News on Thursday, September 13, before the terms of the funding facility were finalised. Panic spread rapidly when depositors became aware that Northern

<sup>48</sup> Marvyn King (Speech at the Northern Ireland Chamber of Commerce and Industry, Belfast 2007) <<http://www.bankofengland.co.uk/publications/speeches/2007/speech324.pdf>> accessed 5 August 2009.

<sup>49</sup> *Ibid.*

<sup>50</sup> *Ibid.*

Rock would collapse and that the deposit insurance scheme did not guarantee 100% of deposits above £2,000. The morning after, a bank run started at the four only branches of the bank in London.<sup>51</sup>

Convinced that the only way to halt the run was to provide a Government guarantee, the Chancellor of Exchequer announced on Monday, September 17 that the Bank of England would supply arrangements that would guarantee all the existing deposits in Northern Rock if necessary. The run was consequently stopped, but the hesitation in providing the emergency liquidity support and the delay in its announcement raised criticism from both the public and Parliament.

### ***Depositor Protection, FSCS and Moral Hazard***

The Northern Rock crisis also drew attention to the weaknesses in the legal framework for dealing with failing banks, especially with respect to the depositor protection scheme. If the bank collapsed, depositors would not receive their deposits in full beyond £2,000 and they would have had to wait months before their deposits were, even in part, reimbursed.

Furthermore, the FSA indicated that the FSCS was not designed to compensate 100% of the deposits, since it would encourage depositors not to monitor banks and to favour institutions which offer high rates of interest. This would place a moral hazard problem, for the depositor would rely on the insurance in case of bank failure, without suffering any adverse consequence. According to Callum McCarthy, “consumer responsibility is therefore vital to the effectiveness of financial markets.”<sup>52</sup>

Nevertheless, the co-insurance terms of FSCS after September 2007 proved to be inefficient in providing stability to the system and to stop a bank run. As the Treasury Committee report concludes, it is unrealistic to consider that retail depositors would have the means, time and ability to assess the financial strength of an institution through the examination of publicly available information about the company. Therefore, rather than contributing to financial stability, co-insurance directly undermines it.<sup>53</sup>

### ***The Special Resolution Regime***

In response to the continued pressure on the financial market since 2007, the Banking Act 2009 introduced a Special Resolution Regime (SRR) with the objectives of enhancing financial stability, protecting depositors, and minimising the effects of a bank failure. It consists of a regulatory insolvency arrangement which comprises of a set of tools that allow the tripartite authorities to deal more accurately with failing banks and building societies. In addition, the Act modified the FSCS, aiming at compensating depositors of defaulted firms promptly.<sup>54</sup>

Following the features of the Federal Deposit Insurance Compensation (FDIC) of the US, the SRR allows the authorities to transfer all or part of a failing bank to a private sector purchaser, a ‘bridge bank’ (i.e. a subsidiary of the Bank of England) pending a future sale or place a bank into a temporary public ownership. Hence, it aims at maintaining the services provided by the firm without severe disruption.<sup>55</sup>

<sup>51</sup> House of Common Treasury Committee (n 24).

<sup>52</sup> *Ibid.*

<sup>53</sup> *Ibid.*

<sup>54</sup> FSA, *Banking and Compensation Reform* (policy statement) (London July 2009).



In addition, SRR enables the tripartite authorities to place a bank into the Bank Insolvency Procedure, which is designed to allow rapid payments to FSCS insured depositors. According to the FSA, the scheme aims at delivering payouts to depositors between 7 to 20 working days by 2010.<sup>55</sup>

As a result, the rules introduced by the SRR enable an orderly resolution of banks in default while protecting retail consumer's interests and financial stability at the same time. However, since the new scheme furnishes the financial industry with an additional safety net arrangement for failing financial institutions, its effects on moral hazard incentives are still to be seen.

## Conclusions

In the financial system, the lender of last resort function of central banks and deposit insurance schemes are the foremost sources of moral hazard. Both create opposite incentives for regulated individuals, whom, protected by these regulatory tools, may threaten the very system that they seek to support in the pursuit of higher profits.

The lender of last resort facility tends to create a dilemma for central banks and financial regulators. While supplying emergency liquidity assistance for banks in distress to maintain stability, this also allows the same institutions to take excessive risk and rely upon the safety net in case of failure, thence increasing the probability of crisis. By injecting liquidity only to the system rather than to individual institutions, central banks enable interbank markets to allocate the resources more efficiently to solvent firms, albeit temporary illiquid. As Thornton and Bagehot suggested, the role of central bank in this regards is to protect the stability of the financial system and to prevent future crises, which assumes that imprudent institutions must be allowed to fail.

One regulatory strategy against moral hazard has consisted of adjusting the incentives of decision makers with regulatory objectives. On the one hand, it has been suggested to increase the regulatory capital for banks, since it places significant incentives for firm owners to invest more carefully and to monitor managers' behaviour. This also constitutes an important measure for regulatory agencies to deal with large banks that operate not only in the banking business, but also in insurance and investment markets, which could lead to higher systemic risk in case of company's failure. On the other hand, after the recent credit crunch, remuneration structures of bank's managers have been a growing concern within the financial regulation and it has drawn the attention of authorities in the search for innovative regulatory tools to prevent future crises.

In implementing these strategies, financial authorities are constantly requiring a broader consensus with regulated firms and regulatory agencies at international level. This is not only to create a level playing field among financial centres, but also to enhance financial supervision.

The bank run caused by Northern Rock in 2007 demonstrates that the co-insurance system of depositor protection undermines the stability of the financial system rather than contributing to the effectiveness of the market. Even though deposit insurance tends to place adverse incentives on consumers, it is delusive to assume that retail depositors are

<sup>55</sup> Bank of England, *Special Resolution Regime* (2009)

<[http://www.bankofengland.co.uk/financialstability/role/risk\\_reduction/srr/](http://www.bankofengland.co.uk/financialstability/role/risk_reduction/srr/)> accessed 23 August 2009

<sup>56</sup> Financial Services Authority (n 55).

able to accurately appraise financial products and institutions, as well as evaluate all the information available to the market. Therefore, financial authorities are demanded to supervise on behalf of depositors, certainly considering the ageing population that increasingly depends upon their savings.

The Northern Rock crisis also indicated that, despite moral hazard considerations, once a liquidity crisis has departed, central banks are required to supply the financial system with prompt corrective actions in order to avoid contagion and maintain confidence in the system. As the bank run on Overend, Gurney & Co. in the 19th century proved, the hesitation of the Bank of England in providing emergency liquidity support to banks may deepen the crisis.

The introduction of the SRR seeks to provide a new approach to failing financial institutions through an orderly process of ownership transfer, as well as to protect depositors with a scheme that compensates their losses quickly. This regime therefore can reduce moral hazard, for it provides the tripartite authorities with a credible option, as Thornton suggested, letting the bank's owners and managers suffer the natural consequences of their fault.



# WHY THE CRIMINAL LAW PRINCIPLE OF CORRESPONDENCE SHOULD PREVAIL OVER THE PRAGMATISM OF CONSTRUCTED LIABILITY\*

DOMINIC J J BRIGHT  
KING'S COLLEGE LONDON  
DOMINIC.BRIGHT@KCL.AC.UK

**ABSTRACT** Why is it significantly objectionable to legal and moral reasoning that constructed liability, a creature of pragmatism, determines so many criminal cases? A defendant should only be prosecuted for resulting harm that was intended and in fact caused: a correspondence of *mens rea* and *actus reus*. The principle of fair labeling thus shares mutuality with, and is parasitic on, correspondence, in that we can place confidence in saying that a convicted offender has been treated justly by a criminal court when there has been commitment to the correspondence principle. This paper begins by adopting Ashworth's statement on the rationales motivating what can be identified as the principle of correspondence and (so) also the principle of fair labeling. The result of the argument maintained here is to launch an independent objection to constructive liability. That is, the form of legal adjudication that sets aside the principle of correspondence as imperative and instead constructs liability. Logically pursuing this argument involves taking issue with the description of constructive liability as a 'principle' of criminal law. In so doing, this paper wholly rejects the proposal that constructive liability is at all rooted in principle and seeks to oust it as a doctrine, or tool, of pragmatism. Considering the rich conception of full-blooded justice that is available to us, the gap in legal justification thrown up by constructive liability is an intolerable anomaly. Hence, this is as much a philosophical treatment of law-making and legal adjudication as it is a criticism specific to criminal law cases.

\* Sincere thanks are owing to Professor Andrew Ashworth, Fellow of All Souls College, Oxford University, who very kindly set aside time to clarify (via telephone) certain aspects of his statements on the principles of correspondence and fair labelling.

## 1. What Fundamental Rationales Motivate the Principle of Correspondence?

Ashworth has identified several principles of criminal law that should be understood as axiomatic to criminal law making and adjudication.<sup>1</sup> The focus in this paper is specifically on two of those principles of criminal law as outlined by Ashworth, which, crucially, should be present in order to establish criminal liability<sup>2</sup> – correspondence and fair labeling. The former is often described by reference to comparable elements of *actus reus* and *mens rea* for the unlawful act: “it should [be] established that the defendant’s intention, knowledge, or recklessness related to the proscribed harm.”<sup>3</sup> This relationship is a fundamental one and should be conceived as such by criminal law theorists and the courts alike – with the result that criminal law at all times stewards and insists on upholding the correspondence principle. That principle requires that before criminal liability is imposed for a given offence, there is a correlation, or relative correspondence, between the requisite *mens rea* and *actus reus*. It follows that if the fault element of the unlawful act – used to gauge the culpability of the defendant – does not correspond with the unlawful act, the defendant should not be prosecuted under the said charge or indictment. It is wrong that a defendant is prosecuted (and many prosecutions do proceed on this basis) by way of constructing his criminal liability.

It is significant that the principles of correspondence and fair labeling are not completely distinct from one another. The latter follows from, and is derived in, the former. Indeed, it is hard to completely detach the principle of fair labeling from that of the principle of correspondence: fair labeling may be achieved through a legitimate correspondence of the culpable mind for the performed unlawful act. Therefore, it becomes possible that a happy coexistence or mutuality is implied and operating between *mens rea* and *actus reus*. In this way, fair labeling derives from the principle of correspondence. If the defendant were to be treated in such a way as to give effect to the principle of correspondence, then by default, the obligation of the courts to fairly label a defendant facing a criminal charge would be discharged. At this juncture, we can also set out the parameters of what is meant throughout this paper by reference to ‘justice.’ The intended sense of justice that is adopted here is what can best be understood as a “full-blooded”<sup>4</sup> conception of justice, since that is surely the only meaningful way in which we speak and write about justice. Care will be taken, in later sections, to difference this full-blooded conception of justice from the more puny (though still worthwhile) sense of justice that is captured by tenets of Natural Justice. Nevertheless, we should be ambitious for criminal law. We should adopt and advance the most robust conception of justice available to us.

However, what is gained through the application of the principle of fair labeling? The principle of fair labeling means:

to see that widely felt distinctions between different kinds of offences and degrees of wrongdoing are respected and signaled by the law, and that offences are subdivided and labeled so as to represent fairly the nature and magnitude of the law-breaking.<sup>5</sup>

<sup>1</sup> Andrew Ashworth, *Principles of Criminal Law* (5th ed. Oxford University Press, Oxford 2006).

<sup>2</sup> It should be stated this paper is not concerned with issues concerning strict liability, whereby expediency and pragmatism prevail over correspondence on grounds of procedural efficiency and ease of regulation.

<sup>3</sup> Andrew Ashworth, *Principles of Criminal Law* (5th ed. Oxford University Press, Oxford 2006) p.87.

<sup>4</sup> Leslie Green, “General Jurisprudence: A 25th Anniversary Essay” (2006) 25(4) OJLS 576. See, also, Leslie Green, “The Germ of Justice” (2006) p.10: “And our question is not whether law application is connected with ‘justice’ understood as the administration of law; it is whether it is connected with justice in some full-blooded normative sense of the term.”

Therefore, not only is the essence of the wrong captured, but also the relative seriousness of the wrongdoing is communicated – fundamental aims of the criminal law. Additionally, if defendants were to be accurately labeled, this may go some distance towards gearing criminal adjudication in the courts to be a more effective deterrent through the creation of a morally defensible social stigma. This would have particular application in the context of conviction for the more serious or heinous of crimes – those higher up the ladder in the criminal calendar.

After qualification of the principles of correspondence and fair labeling, and their virtue when practically applied, it becomes necessary to examine the rationales which underpin them. It seems a missed opportunity that Ashworth did not go so far as to highlight the primary attraction of closely following the principle of correspondence. This is safeguarding the liberty of those most vulnerable – the defendant facing criminal charges and thus potentially subsequent loss of liberty. This becomes apparent when considering the multi-faceted consequences criminal conviction entails. The first consequence may be obvious. The most severe punishment a criminal court (in the UK) can bestow is deprivation of liberty in the form of a custodial sentence. The second is especially prominent in cases where liberty is not lost, or if it is not lost, conceded for a relatively short length of time. This may be termed ‘social stigma’ and, in many cases, is a most resonating aspect of criminal conviction. The ramifications become clear when considering the most heinous crime of all offences – murder. A conviction for murder may give the impression that a defendant (indeed, all defendants) intended and in fact caused death through their actions. As we know, however, this is not invariably the case. An intention to cause ‘really serious’<sup>6</sup> or grievous bodily harm is sufficient *mens rea* for a conviction of murder – an especially striking paradigm of constructed liability.

A defendant’s actions may have caused an unlawful unpalatable consequence (death) but if he did not intend that consequence (and it is not part of the Crown’s case that he did), should he really be charged and labeled as a generic murderer? It is suggested we have good cause to separate (and so delineate) different kinds of homicide that currently fall into the generic category of ‘murder.’ It is precisely this concern for more precise charging and fair labeling that motivated the Law Commission’s Report on Murder, Manslaughter and Infanticide<sup>7</sup> which, in the event, was not adopted by the government. Thus, questions of fair labeling also arise when considering an alternative *mens rea*, which will suffice for the imposition of a criminal conviction.

As will be developed in due course, this route to prosecution is not consistent with the best available conception of justice – full-blooded justice. Liability in criminal cases should, instead, be construed. That is, liability should emerge and take shape if and when there is correspondence between a guilty mind and a guilty act. There is a significant difference between construing liability and constructing liability. The method of construing – for example, an Act of Parliament, an authority, or a finding of liability – can involve purposive interpretations where those can be fairly read in or borne out. However, constructing liability is a different proposition (of legal reasoning) entirely. It involves implying or

<sup>5</sup> Andrew Ashworth, *Principles of Criminal Law* (5th ed. Oxford University Press, Oxford 2006) p.88.

<sup>6</sup> The test for the alternative prosecutorial limb for murder, grievous bodily harm, was formulated by the House of Lords in *DPP v Smith* [1961] A.C. 290, p.334.



reading in that which could not already have been found or derived (say, in the available authorities or legislation). In the criminal context a special significance attaches to this, given the adversarial nature of criminal proceedings. In overview, the background that colours and circumscribes the nature of the criminal trial includes: the need for criminal law to be stated clearly, in advance of prosecution; that the prosecution must establish their case to the satisfaction of the jury so that those jurors can be sure in the absence of reasonable doubt; and that any ambiguity or doubt is to be resolved in a defendant's favour, or at least not against his favour. These pre-conditions of the criminal trial put lie to the pragmatism behind constructive liability. Defendants on trial should be recognised as enjoying latitude to take advantage of the adversarial process where criminal law is silent or criminal law's net has not been cast widely enough. This is to say no more than that the Crown should be put to proof. Constructive liability is, it follows, treated here as a form of pragmatism that betrays a kind of intellectual dishonesty. This is intolerable to the clarity and rationality that legal and moral reasoning should strive to espouse, with all the greater fidelity in the criminal courts.

What then does the correspondence principle involve? Or, put another way, what does correspondence between *actus reus* and *mens rea* demand of us? At the most basic level, the principle of correspondence is premised on law having distinguished between various 'families of criminal offences',<sup>8</sup> adopting the language of categories Gardner deploys. No objection is taken here with Ashworth's conception of the principle of correspondence, in terms of the valuable work that principle does to ensure that criminal law justly disposes of individual defendants. However; as will shortly be outlined in the next section, issue is taken with Ashworth's understanding of constructive liability as a principle or counter-principle. There is good reason to think that our descriptions of constructive liability should be divested of any reference to principle – it is better recognised as a doctrine rooted in, and exhausted by, pragmatism alone.

This understanding is crucial not least because the correspondence principle is routinely neglected by the criminal law, which should be understood as an oversight in legal and moral reasoning. This is for the reason, as will be articulated and developed here, that correspondence is a principle that enjoys a resounding, intrinsic and morally intuitive appeal. Why, despite this, has so much academic commentary attached to the significance and place of the principles of correspondence and fair labeling? One (obvious) reason is that commentators sceptical of the correspondence principle, at least insofar as the principle is conceived as an absolute, non-derogable one, have failed to grasp the essence and importance of the principle.

## **2. What Is Constructive Liability, and Can It Be Reconciled with the Correspondence Principle?**

Constructive liability is referred to by Wilson as an 'ordinary principle'<sup>9</sup> of the criminal law. Ashworth had similarly characterised the correspondence principle as one of several 'principles relating to the conditions of liability'.<sup>10</sup> When pressed on the point, however, Ashworth was prepared to concede that the phenomenon of constructive liability is really

<sup>7</sup> Law Com. No. 304, *Report on Murder, Manslaughter and Infanticide* (2006).

<sup>8</sup> John Gardner, "Criminal Law Theory: Doctrines of the General Part" (2003) L.Q.R. 161.

<sup>9</sup> William Wilson, *The Structure of Criminal Homicide* (2006) Crim. L.R. 924.

<sup>10</sup> Andrew Ashworth, *Principles of Criminal Law* (5th ed. Oxford University Press, Oxford 2006) p.85.

not best understood as a principle of criminal law at all, in that Ashworth clarified or else amended his statement on constructive liability by recognising it as a “counter-principle”<sup>11</sup> to that of the principle of correspondence. It can nevertheless be seen as unfortunate that the language even of counter-principle still imports reference to principle at all. The argument unfolding here is that constructive liability is a form of pragmatism that is in direct conflict with the correspondence principle. There can be no reconciliation between them; in that constructive liability undermines, or else reverses adherence to, the correspondence principle.

We can simply treat this as a difference in emphasis between that which this paper seeks to achieve, and that which Ashworth has taken. The point is that Wilson, too, recognises that the correspondence principle has been plumbed into the corpus of the criminal law. If anything, the quality of ordinariness enjoyed by the correspondence principle only serves to underscore its recognition and function – by criminal law theorists, at least. Wilson then continues to usefully qualify the operation of the principle, or counter-principle, of constructed liability; whereby an offence to which a defendant is culpable is pushed through a higher threshold of wrongdoing.<sup>12</sup> Thus an unlawful act, in the form of a trespass to the person perpetrated by a defendant, may be constructed to form the basis of indictment for manslaughter. However, the counter-principle of constructed liability does not accord the Law Commission with its aim which has been stated as, “[giving] effect to the ladder principle, according to which there should be clear and robust differences between offences of different degrees of gravity.”<sup>13</sup> The rungs of any such ladder principle quickly splinter and give way when correspondence between *actus reus* and *mens rea* is neglected.

So, fair labeling would be served if there were clear and robust differences implied and maintained between offences, reflecting the different degrees of severity. However, the ambition to render effective both the ladder principle and the principle of fair labeling – if not identical principles, so closely akin that there is nothing between them – is all too frequently undermined via the construction of criminal liability. The ladder principle is almost abandoned, especially when considering the way in which the courts routinely construct liability on the basis of offences against the person. Indeed, the judiciary does not generally even acknowledge a need to justify departure from the correspondence principle. This is despite the stringent justification that should be required, given the several pre-conditions of the adversarial criminal trial, before the principle of correspondence is to be neglected or set aside. A prominent issue now arises. The foremost proponent of the principle of correspondence, Ashworth, deems that it may be possible to depart from the principle of correspondence, albeit with ‘adequate’ justification.<sup>14</sup> This is surely an unnecessary concession to pragmatism and expediency at the expense of the demands of full-blooded justice, for reasons that will be developed.

### **3. The Special Nature of Criminal Law Adjudication – The imperative of stringent justification**

The potential deprivation of liberty, combined with the negative social label attached to a

<sup>11</sup> Telephone conversation with Professor Ashworth (29/01/2009).

<sup>12</sup> See, also, Leslie Green, “*General Jurisprudence: A 25th Anniversary Essay*” (2006) 25(4) OJLS 565.

<sup>13</sup> Law Com. No. 177, “*A New Homicide Act for England and Wales?: A Consultation Paper*” (2006) 3.183.

<sup>14</sup> Telephone conversation with Professor Ashworth (29/01/2009).

criminal conviction, should always require justification before sentence is passed. Furthermore, this justification is necessarily of a special kind of order, in that there must be satisfaction of two requirements if it is to be morally defensible. It must be both stringent and adequate. Stringent justification is a mandatory requirement of moral and legal reasoning when liberty is to be curtailed: a defendant must be informed why he is being punished and, equally as important, the wrong should be comprehensible so as to demarcate the scope of the criminal law (and possibly serve as a deterrent).

Stringent justification also performs a potential safeguarding function (for example, against the unwarranted withdrawal of liberty) which, when combined with a general requirement for adequate justification, shows an effective concern for individual defendants. Adducing adequate justification for the purposes of imposing punishment must include demonstrating (as part of the prosecution's case) that the requisite *mens rea* is correlative to the *actus reus*. It follows that the second limb of the *mens rea* which may suffice for murder – an intention to cause grievous bodily harm – should not furnish adequate justification to the prosecution for the imposition of a murder conviction.

This may seem a radical thought that goes against the grain of, and may wholly recant, what many criminal law theorists and those professionally concerned with the criminal courts would view as a received truth – that an intention to ‘really seriously’<sup>15</sup> injure is indeed sufficient for a murder conviction. However, surely this alternative limb of the *mens rea* for murder lumps together those individuals who ought properly to be labelled and convicted as murderers – having the requisite intention to kill, to unlawfully extinguish life – with those individuals falling into this category on the basis of a result not of their design. This is consistent with saying that an individual should be held accountable to law for having set into train a series of events, by way of certain morally impugned conduct. However, it is arguable that more subtle, defensible distinctions should be implied between those convicted of unlawfully taking human life. Again, it should be stressed that this is a requirement of the fair labeling of defendants which should invariably trump pragmatic prosecutions.

The thrust of the argument maintained here is that it is intolerable that English law shows only partial commitment to the correspondence principle. Criminal indictments and prosecutions should locate the *actus reus* and identify the full, correlative *mens rea* in respect exactly of that conduct. Criticism of how the normative interaction between these two is routinely set aside (and can best be exemplified when considering the ladder of non-fatal, non-sexual, offences against the person. Sections 47 and 20 of the Offences Against the Person Act 1861), are all so-called half *mens rea* offences. All that all that is required, for both sections, is culpability only of a common assault and not the harm done. In practical terms, this has the regrettable – and, it is suggested here, intolerable – effect of giving the prosecution another bite of the cherry.

Thus, a prosecution on the basis of the unlawful act of malicious wounding or infliction of grievous bodily harm<sup>16</sup> may be substituted for an assault occasioning actual bodily harm.<sup>17</sup> This is pioneering of the prosecution but wholly unjustified as concerns principles of criminal law: a charge for section 47 requires proof of an assault while a charge of

<sup>15</sup> *DPP v Smith* [1961] A.C. 290, p.334.

<sup>16</sup> Offences Against the Person Act 1861, s.20.

<sup>17</sup> Offences Against the Person Act 1861, s.47.

<sup>18</sup> *R v Dica* [2004] Q.B. 1257.

section 20 does not. *R v Dica*<sup>18</sup> demonstrated that grievous bodily harm could be inflicted through the act of consensual sex and the subsequent transmission of a terminal disease, namely the Human Immunodeficiency Virus (HIV). Thus, the ‘justification’ for the reduction of a criminal charge of section 20, where a defendant is not found guilty by the jury, to the lesser ‘included’ offence (section 47) is vitally flawed – if it were an ‘included’ offence, proof of an assault would be required. Thus, the depth and extent to which pragmatism has infected the criminal justice system must now be apparent.

It is, however, virtuous that the mens rea for the non-fatal and non-sexual offences against the person requires that the defendant is shown to have intended the unlawful harm caused, or foreseen the risk of ‘some harm’ being perpetrated. This is known as *Cunningham* recklessness<sup>19</sup> and is subjective, being assessed on an evaluation of the defendant’s awareness to consequential harm resulting from his actions. This is distinct from the objective test of the reasonable person, known as *Caldwell* recklessness,<sup>20</sup> which has gross potential to permit injustice via a construction of liability. The morality underlying *Cunningham* recklessness is a partial adherence to the principle of correspondence – “relevant culpability”<sup>21</sup> of the defendant. However, *Cunningham* recklessness is not utopian, and is still a diluted form of pragmatism. A foresight of some harm should not render the defendant culpable for unforeseen injurious consequences that stemmed from his actions.

It follows that voluntary intoxication is another expanse of criminal law pervaded by pragmatism. This is because if a defendant were to become intoxicated voluntarily, he would be deemed to have been reckless in law, and therefore worthy to be subject to a construction of criminal liability. However, the rationale underpinning the construction of his criminal liability is fundamentally flawed – reckless in law is distinct and irreconcilable to a general interpretation. This is because in law reckless is understood to mean a defendant foresaw the risk of a consequence of an action and subsequently went on to take it regardless. However, it would be imprudent to regard the majority of defendants who became voluntarily intoxicated to have foreseen a risk that they would commit unlawful activity. Thus, the rationale endorsing the application of prosecutorial construction of liability is morally unjustifiable and as such, should not be permitted.

#### **4. Closely Following the Principle of Correspondence Brings the Criminal Law Closer to Encompassing Full-Blooded Justice**

Several morally and legally significant pre-conditions of criminal proceedings should be respected if full-blooded justice is to be served. One such pre-condition is that there should be minimum application of the criminal law. Another is that the prosecution must state its case, and prove that case beyond all reasonable doubt. A third is that, in the event of any doubt, a defendant should enjoy the benefit of the doubt. Furthermore, where criminal law is applied it should be appropriately justified, not only legally but also morally, in order for the criminal law to reflect the full panoply of these so-called pre-conditions of the criminal trial. In one sense, minimum application of the criminal law may indicate an unwillingness to invoke criminal convictions upon defendants. Or, put differently, minimum criminalisation readily speaks to the thought that law should be slow and

<sup>19</sup> *R v Cunningham* (CA) [1957] 2 Q.B. 396.

<sup>20</sup> *R v Caldwell* (HL) [1982] A.C. 341.

<sup>21</sup> CMV Clarkson, *Understanding Criminal Law* (4th ed. Sweet & Maxwell, London 2005) p.15.

reluctant to deprive individuals of their liberty. Law enjoys the quality of being gentle. The advantages in this sense may appear obvious: the State is not excessively subjected to a financial burden which is without due cause; and when criminal convictions are imposed they may have sufficient impact and resonance in order form an effective deterrent to subsequent unlawful (and thus criminal) behaviour.

In a more implicit sense, minimum application of the criminal law may well lend itself to the fair labelling of defendants, such that the conviction a defendant may receive is akin to his perpetrated unlawful act – and thus, not greater than. Another such pre-condition of a criminal trial is the principle that if there is any ambiguity during trial, it is resolved in favour of the defendant, or at least not against his favour. This is also consistent with the underlying principle that a defendant is innocent until proven guilty. If the criminal law is to curtail the liberty of its citizens, it must be both lawful and justified upon morally defensible rationale. Thus it may be argued that the construction of criminal liability is contrary to the rule of law, to which we are all subject.

We need not shrug our shoulders as to what is meant by ‘justice’ in terms of how the criminal courts should dispense of punishment. There is a concept of justice, in the way of a canonical formula - the origins of which can be located in Roman ideas of justice: giving what is due, to whom it is due, from whom it is due.<sup>22</sup> This captures the idea that justice can be conceived in terms of *allocations*. It follows that to achieve this definition of justice the defendant must receive ‘what is due.’ A construction of criminal liability does not achieve this outcome because the *mens rea* of the defendant has undergone fabrication in an attempt to render conviction sufficiently justified – but not commensurately so, due to the absence of a state of culpability.

It is central to this paper that the conception of full-blooded justice, which can be identified as a so-called ‘thick’ conception, is considered distinct from justice otherwise confined to principles of Natural Justice. It is possible to highlight this distinction: sentencing passed may be unjust to the defendant although the procedural principles of Natural Justice were strictly adhered to. It now becomes necessary to outline the principles of Natural Justice, which are that a defendant has a right to a fair hearing by an independent and unbiased arbitrator. Thus, even if a defendant benefits from a fair hearing, for example, a construction of his liability will render possible a conviction of the indictment laid before him even though he may not be culpable of the offence – because the principle of correspondence is neglected and hence there is an absence of a thick conception of justice.

## **5. Actus Reus and Mens Rea Must Coincide in Time for Criminal Liability**

To establish criminal liability, not only should there be a correspondence of *mens rea* and *actus reus*, but these elements should also coincide in time. This principle upon the establishment of criminal liability should be so elementary so as to be irrelevant to our discussion. If a defendant, after rowing with his wife, intended her death and years later through his actions her death was caused inadvertently *de facto*, should he be liable to a charge of murder? The answer is morally coherent – no, he should not, however distasteful his fantasy. Despairingly it is not so clear in law. Even this cardinal principle has been

<sup>22</sup> Leslie Green, "General Jurisprudence: A 25th Anniversary Essay" (2006) 25(4) *OJLS* 565.



warped and distorted in order to permit a prosecution to construct criminal liability via two further distinct principles. The first of which has come to be known as continuing acts, whereby the *actus reus* is argued to have been a continuing act to which a later *mens rea* can therefore coincide. It was on this pretext that a defendant was successfully prosecuted for the unlawful act of driving his car onto a policeman's foot unwittingly, albeit having realised the gravity of his act subsequently.<sup>23</sup>

The second principle invented by the judiciary, in which coincidence of *actus reus* and *mens rea* may coincide, is where there has been one continuing transaction. Thus, during the period in which the unlawful act was taking place, if the defendant formed the necessary *mens rea*, criminal liability of the defendant can be constructed.<sup>24</sup> However, these principles, via which it is possible to construct criminal liability (through a somewhat artificial coincidence of *mens rea* and *actus reus*) are not to be regarded as incompatible with the principle of correspondence. While this approach to constructing liability is fallible on grounds of its artificiality, it should however be recognised. It does not pose an intolerable threat to, and so distract from, the discharge of full-blooded justice which this paper has argued should be the primary concern of the criminal law.

## 6. Why Horder and Clarkson Cannot Defensibly Justify Any Departure From the Principle of Correspondence via the Creation of Exceptions

Before continuing, it must be made clear that there is no issue to be taken with the principle of transferred malice.<sup>25</sup> This is significant to note because transferred malice is a principle used as a tool of pragmatism to bring about the construction of a criminal conviction. It is easily exemplified by the paradigm in which A intends to kill B. However, A shoots and in fact kills C. A has carried out the *actus reus* of murder: unlawfully killing a reasonable person who is in being under the King's Peace.<sup>26</sup> A does have an intention to kill, which is adequate *mens rea* for murder despite his actions resulting in the death of another citizen who is not the intended target. Thus, the practical application of transferred malice may be put simply. The defendant's *mens rea* is transferred upon the victim de facto – in this case, C.

It becomes morally defensible to construct the defendant's liability, within the limited scope of the principle of transferred malice, because there is correspondence between the *mens rea* and *actus reus*. Therefore, the defendant can justifiably be rendered culpable for his unlawful action, which was the “operating and substantial”<sup>27</sup> cause of the death of C. It is important, however, that there is a justifiable scope limiting the range in which a defendant's malice can be transferred upon the victim in fact. Thus, an intention to cause grievous bodily harm to B cannot be used as a basis to construct a defendant's liability for the murder of C.<sup>28</sup> This confinement upon the application of the principle of transferred malice is necessary, rendering the current law's application of the principle morally defensible, as well as practical.

<sup>23</sup> *Fagin v Metropolitan Police Commissioner* [1969] 1 Q.B. 439.

<sup>24</sup> *R v Thabo Meli* [1954] 1 W.L.R. 228.

<sup>25</sup> See, *R v Latimer* (1886) Q.B.D. 359.

<sup>26</sup> See, also, Clarkson and Keating, *Criminal Law Text and Materials* (6th ed. Sweet and Maxwell, London 2007) p.663.

<sup>27</sup> *R v Smith (Thomas Joseph)* [1959] 2 Q.B. 35. See, also, ‘contribute significantly’ *R v Pagett* [1983] Cr. App. R. 279.

<sup>28</sup> *Attorney-General's Reference (No.3 of 1994)* [1998] A.C. 245.



## 7. Luck and the University Fundraiser – A confused and ultimately unworkable analogy when lifted into the criminal context

Clarkson utilises a paradigm, first enunciated by Horder, in an attempt to explain why it is that a defendant “makes his own luck.”<sup>29</sup> If a university fundraiser in soliciting for donations of £100 from alumni and in the event an alumnus in fact donates £500 to the university, this is not *pure luck*. This is because the university fundraiser has actively engineered a scenario in which it would have been foreseeable, albeit possibly not probable. However, the paradigm is continued thus. If an unknown alumnus donated £500, Horder argues this would be pure luck; for the reason that no steps had been taken by the university fundraiser or put into chain. However, this analogy fails to delineate the pre-conditions peculiarly unique to a criminal trial discussed above – minimum application of the criminal law; ambiguity is resolved in favour of defendant, or at least not against his favour, and so on. Thus, the scenario is entirely different in the paradigm example of the university fundraiser, so much so, that it is hard (if not, in fact, impossible) to reconcile with that of criminal proceedings.

It will be argued that, save the principle of transferred malice, there is no morally defensible justification for the departure from the principle of correspondence and the subsequent imposition, via construction, of criminal liability. A primary but substantial step in evidencing this proposition would be the demolition of the main school of opponents or sceptics to a mandatory implementation of the principle of correspondence. If this task of demolition is achieved; it does not, of itself, establish or propagate the correspondence principle. Rather, it would do significant work to this effect.

## 8. The Importance of Preserving the Notion of ‘Pure Bad Luck’

Clarkson is of the same opinion as Horder: while it is possible, intellectually, for a defendant to be the victim of pure bad luck, his moral inculcation in a course of events leading to a dire result is an efficient basis on which to construct his liability. The premise to this argument is that the defendant has pro-actively engineered circumstances in which it is not *unforeseeable*, albeit perhaps not foreseeable, that an unlawful act may come to fruition in a certain way. Clarkson and Horder are able to take this view due to the shared belief that a defendant makes his own luck. However, the practical significance and real worth of this argument is that it is an attempt at a justification for prosecutorial construction of criminal liability. Thus, “those who intentionally attack others are morally responsible for and so fully legally accountable for the consequence of so doing whether or not such consequences were foreseen.”<sup>30</sup> However, a concession may be extracted from Wilson, who has acknowledged that this is the “uncomplicated view.”<sup>31</sup> This leads one to ask: should the criminal law be concerned with the uncomplicated view – the *easiest* way to dispose of criminal defendants – or the most morally defensible, and therefore just, view?

When the defendant alleges pure bad luck and is considered to have taken proper care in accordance with the situational circumstances, pure bad luck may in fact have eventuated.<sup>32</sup> The normative position of the defendant should not be anticipated and hived off in this way. The notion of a change in normative position should be recognised as one that is

<sup>29</sup> Jeremy Horder, ‘A Critique of the Correspondence Principle in Criminal Law’ (1995) *Crim. L.R.* 759. See, also, Jeremy Horder, ‘Rethinking Non-Fatal Offences Against the Person’ (1994) 14 *OJLS* 335.

<sup>30</sup> William Wilson, ‘*The Structure of Criminal Homicide*’ (2006) *Crim. L.R.* 475.

<sup>31</sup> *Ibid.*

<sup>32</sup> Telephone conversation with Professor Ashworth (29/01/2009).

suspect (and possibly, a misleading one), providing inadequate justification for the layer of additional liability it imports. “[N]either the purported justification nor the limiting principles sometimes attached to it are convincing reasons for enhanced liability.”<sup>33</sup> This may be substantiated by reference to the law of torts (civil wrongs); whereby an individual is only held liable if holding him to account ordinarily passes the objective test, which is reasonable foreseeability. Additionally, the principle of *novus actus interveniens* (a break in the casual chain) is cognisant of wholly unforeseeable situational circumstances and therefore relieves the defendant of liability for the unexpected harm, which must be deemed independent of his actions. Thus, pure bad luck may indeed be possible: the defendant took reasonable precautions to avert liability and the turn of events does not pass the reasonably foreseeable test.

An implicit criticism of the tool of pragmatism can now be fixed. When constructing a defendant’s liability, it is a confusion between the outrage and contempt we may have for the unlawful act itself, with the special task the courts must be seen to dispense and in fact discharge justice. The outrage of the unlawful act can be dissected into two segments: the forms of the actual harm caused by the perpetration; and the contempt of law required for the manifestation of the act. Thus, due to this multi-faceted infliction of socially repugnant behaviour, it can be comprehended why a prosecution is willing to construct criminal liability. However, this does not *justify* the existence, and application, of the tool of pragmatism itself. Thus, a bereaved husband who seeks revenge for his wife’s murder may choose to act as vigilante, seeking revenge upon her killer, however, this is *understandable* and not *justified*. It should not be merely sufficient for the construction of liability to be “palatable.”<sup>34</sup> Criminal sanctions should be imposed with both disinclination and stringency as well as intelligibly, because it is the highest renunciation available for imposition upon a member of a civilised society. Outrage and contempt, via the social stigma shackled to a criminal conviction, could be reflected with real gravity by a lesser charge (of which a defendant may actually be culpable).

Holder and Clarkson fashion the argument in the following way. When a defendant has crossed a moral threshold, the construction of his criminal liability is thus warranted. To delve deeper and so test this approach, an extension to the argument that a defendant has crossed a moral threshold is in the claim. Due to the illegality of the alleged act, the defendant should be the last person to claim illegality of procedure because of his morally compromised position. Thus, Holder would argue that a construction of criminal liability, a ratcheting-up of indictment to a higher offence (carrying an increasingly severe sentence), is justified on this premise. However, “the idea of crossing a moral threshold should not be strong enough [for the basis a construction of criminal liability].”<sup>35</sup> This is the most defensible statement: a demarcation of where the moral threshold for where a particular charge in fact lies is unquantifiable and as such, should not be a foundation to justify the utilisation of the tool of pragmatism (constructed liability) and subsequent imposition of a criminal conviction.

<sup>33</sup> Andrew Ashworth, ‘A Change of Normative Position: Determining the Contours of Culpability in Criminal Law’ (2008) 11 No. 2 New Crim. L.R. 232.

<sup>34</sup> Law Com. No. 304, *Report on Murder, Manslaughter and Infanticide* (2006) p.479.

<sup>35</sup> Telephone conversation with Professor Ashworth (29/01/2009).

## 9. Conclusions Motivated by the Thesis Idea

Several conclusions further illuminate why the normative position of the morally impugned defendant should only be stringently *construed* – and not *constructed*. A major criticism of criminal law theory raised here was that constructive liability is not best or, it has been argued, even adequately understood as a *principle* or even *counter-principle* at all.

Certainly, close adherence to, and the unrestricted application of, the principle of correspondence – even while permitting no possible exceptions enabling a departure from that principle – does not *guarantee* the discharge of full-blooded justice. That richer conception of justice should be understood as the organising ambition and purpose of criminal law. However, combined with rights arising from English common law, including procedural adherence to the full panoply of principles of Natural Justice, in tandem with Article 6 of the European Convention on Human Rights (ECHR); it brings criminal proceedings *closer* to fulfilling this primary ambition that criminal law proceedings must necessarily have. Put simply: fidelity to the correspondence principle must be understood as a necessary prerequisite or integral part of criminal prosecutions and adjudication. This is imperative for the reason that judicial conformity to the principle of correspondence logically precedes the fair labeling of defendants.

The construction of criminal liability – with the lone exception, as identified above, as the application of the principle of transferred malice – should not be a viable prosecutorial or judicial option. Constructive liability should in no way feature in criminal law-making and adjudication; it is of no fit ideology in the domain of the criminal courts. Constructive liability represents an unjustified departure from the principle of correspondence and thus forms a live basis for injustice, contrary to the special nature and ambitions of criminal adjudication. The pragmatism involved when constructing the criminal liability of a defendant confuses, on the one hand, the contempt and perhaps outrage we may understandably have for an individual's morally impugned conduct with, on the other, the special task that the courts must discharge – full-blooded justice.

This exposes as a sham the argument that, once a defendant has perpetrated unlawful activity, he has crossed a moral threshold. Instead, allied with Ashworth, it was argued here that that approach is, “not strong enough to act as a justification for the departure from the correspondence principle.”<sup>36</sup> Clarkson's contention – “while the defendant's culpability does not correspond with the [unlawful act] there is ‘relevant culpability’<sup>37</sup> – cannot be reconciled (or indeed rendered consistent) with the more morally defensible and sound standpoint championed by Ashworth. Adherence to the principle of correspondence deems necessary the requirement for at least a ‘relative correspondence’<sup>38</sup> of *mens rea* and *actus reus*. The rationales for, and thus, the importance of, criminal law's insistence on adherence to the principle of correspondence (and necessary rejection entirely of constructive liability) were developed and ordered here. The result of the thoroughgoing arguments maintained is that a fresh statement on the correspondence principle has here been advanced. This goes further in point of principle and commitment even than Ashworth, whose adoption of that principle has caused it to be taken seriously.

<sup>36</sup> Telephone conversation with Professor Ashworth (29/01/2009).

<sup>37</sup> CMV Clarkson, *Understanding Criminal Law* (4th ed. Sweet & Maxwell, London 2005) p.15.

<sup>38</sup> Telephone conversation with Professor Ashworth (29/01/2009).

# SELECTIVE ABORTION: SELECTING THE RIGHT RESPONSE

CHRIS COWLAND, LLB  
KING'S COLLEGE LONDON  
CHRISCOWLAND6@HOTMAIL.COM

**ABSTRACT** The purpose of this article is two-fold. The first is to explore the legal and moral justifications in favour of, and criticisms against, the selective abortion of fetuses likely to be afflicted with serious impairments and handicaps. From this investigation it will be posited that the best justification for allowing selective abortion is rooted in the argument from parental interests and the autonomy of the mother. The second is to isolate the issue of the regulation of selective abortion and the appropriate role of the law. Having discredited a total ban it will not be argued that selective abortion should receive minimal or negligible legal scrutiny. Instead the compelling justifications for state involvement will be highlighted. This together with the aforementioned parental interests/autonomy argument will then inform the view that English law currently draws unnecessary divisions between pregnancies that do and do not involve disabled fetuses pre-24 weeks. Consequently a disability ground for abortion before that point need not exist. That this distinction exists will also accentuate the inadequate approach post-24 weeks. Past this point of viability it will be argued that the general interests of the parent should be subordinated to that of the foetus, as is the case in 'normal' pregnancies other than those in which the mother's life or health is seriously endangered. In other words, only where it is truly in the best interests of the foetus to be aborted will it be permitted. Based on these conclusions legal reform will be proposed and scrutinised.

## Introduction

The most controversial ground of the Abortion Act 1967 is s.1(1)(d). It allows abortion at any stage during the pregnancy where there is "a substantial risk that if the child were born it would suffer from physical or mental abnormalities as to be seriously handicapped." Much of the controversy revolves around the lack of definitions for "substantial risk" and "serious handicap" with concerns over the wide discretion afforded to the medical profession. But more fundamentally there is a dispute as to whether selective abortion of defective fetuses is morally and legally justified in the first place. One cannot escape the

fact that the foetus is to be aborted purely because it is suffering from a disability. Therefore, as with much of medical law, the area stirs strong feelings and emotions. It is submitted in Part I of this paper that despite much argument to the contrary there is a justification for selective abortion of a defective foetus under the law. This is so without refusing to acknowledge the worth of all fetuses. The arguments against such a ground, whilst compelling, do not prove the case they set out to. No submission will be made, however, that supports a total deregulation of abortion in this context. Nor does this paper argue that the current law is satisfactory. It will be argued in Part II that it makes unnecessary distinctions between ‘normal’ and disabled fetuses and thus requires amendment. Possible amendments will therefore be suggested and their legal, moral and medical implications examined in conclusion.

## 1. Legal and Moral Justifications For and Against Selective Abortion

In answering the question whether it is legally/morally justified to abort a foetus under s.1(1)(d) of the Abortion Act 1967, we encounter a number of possible justifications and responses in the negative. In 1.1 the unsuccessful arguments for and against shall be explored whilst in Part II the best justification.

### 1.1 Common Justifications for Prohibition – The inherent flaws

One possible justification for selectively aborting the defective foetus is the best interests of the foetus. The so-called ‘foetal interests argument’<sup>1</sup> posits that since the child would suffer from a serious handicap, it would be better to spare them such a life by terminating at the foetal stage. However, many commentators have highlighted this as a wholly inappropriate general justification. Sheldon and Wilkinson contend that “in most cases the resultant child will have a quality of life which although arguably less good than it would have been without impairment, is still positive and still a life worth living.”<sup>2</sup>

It is submitted that this view must be correct. It is difficult to argue that to live with Down’s syndrome or cleft palate would be so unbearable that it would be in the best interests of the foetus to be aborted and spared such a life. It might be said that a child suffering from Down’s syndrome has a less fortunate start to life due to reduced cognitive abilities and a higher risk of conditions such as congenital heart defects. But the simple fact is that a Down’s syndrome child who receives suitable education and care is not at all prevented an enjoyable, fulfilling life full of worthwhile experiences. This is certainly going to be true of many conditions. For example in a recent study by Newcastle University it was “demonstrated that children with cerebral palsy have high levels of happiness not significantly different from those of the general population.”<sup>3</sup> Rosamund Scott concludes that where birth is compatible with having a good or reasonable quality of life then “we cannot say that it is not in the interests of a child who would have... such a life to be born.”<sup>4</sup> However, whilst discredited at this point the foetal interest’s argument does become relevant later in relation to the distinction of pre and post-24 weeks and the

<sup>1</sup> Sheldon and Wilkinson, *Termination Of A Pregnancy for Reason of Foetal Disability: Are There Grounds For a Special Exception in Law?* 9 MLR 85 (2001), 88.

<sup>2</sup> *Ibid.* 90.

<sup>3</sup> The Lancet, Volume 369 Issue 9580, 30th June 2007.

<sup>4</sup> R Scott, *Choosing Between Possible Lives: Law and Ethics of Prenatal and Preimplantation Genetic Diagnosis*, 1st edition (Hart Publishing Oxford 2007), 87.



proposals for legislative amendment set out in Part II of this paper.

One of the criticisms of selective abortion is that it is eugenic in nature. Critics assert that it seeks to allow only the birth of children with desirable characteristics and essentially murders those who lack such attributes or more accurately, who possess undesirable traits. Wolfensberger equates the abortion of handicapped fetuses with infanticide. He contends that it is the moral equivalent of genocide towards infants.<sup>5</sup> However, Peter Byrne points out that this contention “is based on a particular view of the human conceptus to which reasoned alternatives are available.”<sup>6</sup> Essentially it depends on saying that a foetus is equal to that of a child. It is worth exploring the status of the foetus in some detail here as it does have a bearing on the legitimacy of all abortions and is relevant when discussing the current inconsistencies between selective and non-selective terminations in Part II. The assertion that a foetus can be murdered and that selective abortion is the same as infanticide cannot be recognised as correct even if one distinguishes between a legal person with defined legal rights and a moral person. The position of the law is that a foetus is not a *legal* person. Indeed the European Court of Human Rights (ECtHR) has held that it does not have the right to life under Art 2. In *Paton v UK*<sup>7</sup> the Court stated that the unborn child could not be regarded as a ‘person’ protected directly by Article 2 of the Convention. This does not mean however that the foetus is therefore not entitled to any protection. As Scott points out, protection can be provided to foetal life that need not be tantamount to granting the foetus rights, including the right to life.<sup>8</sup> She notes that by not giving the mother a right to abortion this potentially shows how Parliament appreciates the value of the foetus, affording it protection without having to bestow any rights upon it.<sup>9</sup> Indeed, as will be explored later, in the American decision of *Roe v Wade*<sup>10</sup> it was established that the state has a legitimate interest in the preservation of life which becomes compelling after 24 weeks. It is the foundation of the submissions in this article that there is a similar rationale in the UK. However, having just accepted how legal protection afforded to the foetus falls short of any actual legal rights this seems to allow us to deny the claim that, legally, murder of the foetus is possible. There must be a recognised legal right to life before one can be murdered and at present, there is no such right for the foetus. But can it still be convincingly argued, as Wolfensberger contends, that aborting the foetus for handicap is viewed as morally equivalent to killing a child. It is submitted that this argument lacks merit. No assertions shall be made that the foetus, handicapped or otherwise, is worthless or without value. Such a claim would be highly repugnant and it is uncontroversial to submit that the reverse must be true. However, Feinberg’s view that a foetus cannot be a moral person, that it lacks actual personhood and therefore as a matter of morality lacks any right to life is persuasive.<sup>11</sup> It is submitted that it is true that a foetus cannot be viewed as a person and therefore morally cannot be murdered. Feinberg states that whilst the continued development of the foetus does not grant a right to life nor allow us to say that abortion is murder it does, however, provide a claim to life, an increasingly strong *reason*

<sup>5</sup> W Wolfensberger, *The New Genocide of Handicapped and Afflicted People* 3rd (Rev) edition (Syracuse University, New York 2005) 68-69.

<sup>6</sup> P Byrne, *Philosophical and Ethical Problems in Mental Handicap*, 1st edition (Macmillan Houndmills, Basingstoke 2000), 97.

<sup>7</sup> 1980 3 EHRR 408, 413.

<sup>8</sup> Scott (n4) 108.

<sup>9</sup> *Ibid.*

<sup>10</sup> 410 U.S 113 (1973).

<sup>11</sup> J Feinberg, *Freedom and Fulfilment, Philosophical Essays* (Princeton University Press 1994), 55.

<sup>12</sup> *Ibid.*



not to abort.<sup>12</sup> This will become very relevant later when discussing proposals for the direction that the law should take. Indeed it will seriously question the idea that a foetus can be aborted right up until birth if there is substantial risk of serious handicap as currently interpreted. But it does not at this juncture mean that selective abortion is morally wrong and it is submitted that Feinberg's comments deal with the accusation of genocide. He suggests that the closer to actual personhood the greater the case would be for saying that killing a foetus could be wrong "even though it violated no rights of the fetus, even though the fetus was not a moral person, even though the act was in no sense a murder."<sup>13</sup> Two further points should be noted. In accepting the foetus as a moral person it would seem this could lead to the *prima facie* charge that all abortions would be murder, the intentional killing of a person. Also, affording a right to life to the foetus would lead to a serious moral dilemma. One could imagine a situation where the doctor and the law are faced with two competing rights to life. What happens in a hypothetical situation where you can only save one, for example where a mother could only survive with an abortion. The notion of one enshrined, inalienable right to life coming up against another, whilst not unknown to the law, is deeply problematic.

Moving away from the interests of the foetus, criticisms of selective abortion have been made about its impact on the disabled community.<sup>14</sup> The idea that aborting a child because it is handicapped can be seen as identifying the disabled as persons whose lives are not worth living, which is at first compelling. It might appear that allowing a foetus to be aborted because of the inevitable handicap(s) the child will suffer devalues people who do have those conditions. Imagine a couple who decide on an abortion because the child will be afflicted with Down's syndrome. Based on the aforementioned argument that a person with Down's syndrome can still live a fulfilling life, is it not then arbitrary to deny the existence of a person purely because of their possession of that characteristic?<sup>9</sup> It suggests that the couple and society by permitting this abortion are saying to the disabled person, "Whilst we accept that you might live an enjoyable life, on balance, it is still better that you did not exist." But it is submitted that whilst it could indeed be stated that it is surely better that a foetus be free from handicaps and mental/physical impairments it does not follow that this points to an existing individual and says that it would be better that they had not been born. One can say that ideally it would be desirable that a child not suffer from Down's syndrome but this does not mean that the person with Down's syndrome is condemned or viewed with any less worth. The affliction and the actual afflicted person should be regarded as separate. Byrne notes how:

It is right to regard cognitive disability in the abstract, in isolation as negative in value, but it is right to regard the concrete whole who is the child (the person) with the disability as of intrinsic, compelling worth.<sup>15</sup>

Selective abortion aims for the prevention of a potential person who would suffer from some affliction. It is not focused towards a particular individual. A human individual who is suffering from some kind of impairment "cannot lose the moral status that comes with being an individual human being."<sup>16</sup> The thrust of this argument seems correct and it has found favour among other academics. Shakespeare argues that parents faced with the

<sup>13</sup> *Ibid.* 56

<sup>14</sup> M Field, 'Killing "the Handicapped"' (1993) 16 *Harv Women's LJ* 79, 115

<sup>15</sup> Byrne (n6) 102

<sup>16</sup> Byrne (n6) 100

prospect of caring for a disabled child, which could well be more arduous than raising a normal child, have a “right to forgo this future, which does not imply that they do not like, respect and accept disabled people.”<sup>17</sup>

If one accepts that selective abortion does not point towards a person and degrade their existence could it still be argued that it marginalises and segregates disabled people by a decrease in their number?<sup>9</sup> Martha Field contends that one legitimate view is that there could be societal disregard for the lives of those who remain. The contention is that such disregard on the part of wider society would manifest itself in a decrease in the resources and time dedicated to those with disabilities.<sup>18</sup> Although not phrased in such terms by Field, it is submitted that the underlying argument is best viewed as grounded in utilitarianism and that even if the assumptions she makes were true it would still be inappropriate to rely on utility. Consider the following: One might justify outlawing selective abortion because it will bring the greatest happiness to the greatest number. Happiness will persist throughout the disabled community as they will not suffer from a lack of public support or face the threat of marginalisation. Reassurance, strength and enhanced well being shall inevitably follow from being part of a larger afflicted group. Happiness will be experienced by the rest of society as their lives are enriched by increased diversity and variety of persons. Assuming this were indisputable it is submitted that it would still be wrong to require a pregnant woman to go through a pregnancy, to force her to carry a child to term she did not want because of the happiness that would be experienced by the rest of society. It is a sacrifice of autonomy that cannot be justified by an aggregate increase in happiness. But it is submitted that it does not follow that a decrease in the numbers of disabled persons would necessarily lead to loss of support and exclusion from society. Indeed it has been pointed out that the disability rights movement has not met this evidential burden.<sup>19</sup> It has also been highlighted that this will depend on a number of factors and variables, “not the least significant of which is whether the public is alerted in advance to the danger of reduced support.”<sup>20</sup> This could take the form of a continued awareness campaign. A decrease in numbers might even provoke more stringent rights protection for the disabled minority. In addition with a continued commitment to equal opportunities and access to all walks of life for disabled persons then there is no reason that marginalisation would follow despite a fall in their numbers. Peter Byrne also draws our attention to the fact that despite the presence of selective abortion “there will still be very many injured, disabled and damaged people in society who will keep our sense of compassion and justice alive.”<sup>21</sup>

## **1.2 Parental Interests and Personal Autonomy – The most satisfactory answer for selective abortion**

Notwithstanding the fact that there may be something of a ‘taboo element’ in stating that a parent has an interest in whether or not to have a disabled child<sup>22</sup> it is submitted that reproductive autonomy and parental interests are the best justifications for selective abortion. Scott has stated that:

<sup>17</sup> T. Shakespeare, *Disability Rights and Wrongs* (Routledge, London 2006) 97

<sup>18</sup> Field (n14) 117

<sup>19</sup> Buchanan, Brock, Daniels, Wilker, *From Chance to Choice: Genetics and Justice* (Cambridge University Press, Cambridge 2000) 266

<sup>20</sup> *Ibid.* 267

<sup>21</sup> The Lancet (n6) 101

<sup>22</sup> Scott (n4) 90

Since people can decide whether to reproduce and since extensive positive duties are ... entailed in raising a child... prospective parents should have some say about whether they give birth to any child, including a disabled one.<sup>23</sup>

This rightly identifies the duties and responsibilities involved in becoming a parent – that the mother must care and provide for the child. The issue is essentially one of self-determination and the law must respect this reproductive autonomy. At the basic level it does so by allowing access to abortion through the Abortion Act 1967. It is important to note at the outset that in accepting the autonomy of the pregnant woman and her interests this paper is not advocating that there should be neither regulation nor restrictions in relation to selective abortion. This will be explored later. For present purposes it will be shown that the reproductive autonomy of the mother and idea of parental interests dictates that it would be wrong to impose an *outright* ban.

So, accepting the underlying reproductive autonomy of the mother as the starting point, the justification for access rests on the presence of an inconsistency that emerges if we argue that selective abortion should be prohibited in every case. Tom Shakespeare has highlighted such an inconsistency that arises between social and selective abortion if the latter is prohibited.<sup>24</sup> The position submitted in this paper, and supported by Shakespeare, can be best defended by reference to the following theoretical situation. Imagine two women come to a clinic/hospital seeking an abortion in the 12th week of their pregnancy and assume a ban on selective abortion is in force. Suppose the first woman, X, is seeking an abortion under s.1(1)(a) of the Abortion Act. This provides for an abortion where the “continuance of the pregnancy would involve risk, greater than if the pregnancy were terminated, of injury to the physical or mental health of the pregnant woman.” It is often referred to as the ‘social ground’ and is by far the most invoked of the statutory grounds. Now suppose that the second woman, Y, wishes to abort for disability. It is hard to advocate that as between X and Y, who as persons are legally and morally afforded the same rights and autonomy, one need only seek the approval of two doctors (which she will inevitably find) whilst the other has *no choice* but to carry the child to term. This becomes clearer with further details. Suppose that X’s specific reason for seeking an abortion is that she feels she does not have the financial resources to take care of a child or that she feels it would impact upon her career aspirations. Being forced to carry the baby to term could clearly represent a risk to her mental and physical health greater than if the pregnancy were terminated as required by s.1(1)(a) taking into account her actual or reasonably foreseeable environment. A doctor may not believe this is correct and could conscientiously object under s.4(1) of the 1967 Act but it is almost unthinkable that in the UK she would never be able to find two doctors to agree to carry out the procedure. Of course, they might (and in my opinion should) highlight adoption as a viable alternative but ultimately they would surely (rightly) defer to X’s wishes. But suppose Y felt that she would be unable to raise a disabled child because that child would demand more in terms of care and cost than she feels she can cope with and that it would be deleterious to her overall wellbeing. It does seem to be somewhat arbitrary when examined in these terms that she would be prevented from ever having an abortion for this reason. Whilst the second reason is related to the disability of the future child Shakespeare has rightly alluded to the fact that this could certainly be

<sup>23</sup> *Ibid.* 86

<sup>24</sup> Shakespeare (n17) 93-94

regarded as no more eugenic in nature than the sorts of reasons we attributed to X in this scenario.<sup>25</sup> Whilst it might be true that “a woman will often find it difficult to obtain an abortion under this [social] ground after about the sixteenth week,”<sup>26</sup> the fact is that she would not be absolutely barred whereas she would be if selective abortion was prohibited as some wish. To absolutely prevent selective abortion would completely undermine the autonomy of women, adding restrictions to her choice that are not present where the foetus is not impaired. In both cases self-determination is the key.

One might go further and say that the case for allowing the woman autonomy is even greater in the selective abortion context. This is based on the argument that the psychological stress involved in raising an unwanted child is likely to be greater when that unwanted child is likely to suffer a disability.<sup>27</sup> The British Medical Association describes how ‘parents experience deep shock at the loss of what they had believed previously was a normal pregnancy’ often experiencing symptoms of acute grief including “anger, despair, guilt, inadequacy, sleeping and eating difficulties.”<sup>28</sup> If this is the case then instinctively it might seem wrong to force a mother to undergo such feelings of greater intensity. It might also be the case that there are high levels of stress and worry about the financial implications and how the mother is to facilitate the child’s integration into society. The need to adapt a family home, to arrange special schooling and an enhanced desire to shield her child from the outside world could all be real concerns of the woman. If reproductive autonomy is respected where contemplation of the implications is arguably less stressful, where raising the child will be less emotionally and economically burdensome then doesn’t this seem unfair to the woman seeking selective abortion?

In reply it could be argued that in the majority of cases raising a child with cognitive impairments does not in reality equate to an economic or emotional burden any greater than does raising a child who is not endowed with such a condition.<sup>29</sup> Simon Vehmas opposes selective abortion arguing that:

Every child is more or less a burden to her parents. Children without impairments may cause stress to their parents due to problems (e.g. drug and alcohol abuse and eating disorders) which children with cognitive impairments usually do not get involved in.<sup>30</sup>

However, it is submitted that, should this point be conceded, the objection actually shows that it is no more morally wrong to abort an impaired foetus than one who will not suffer from such impairment. In other words it adds to the assertion that there is an inconsistency. It does not support an outright ban on selective abortion. The woman who does not feel ready for the responsibility of raising and caring for a child, of contending with the possibility of drug and alcohol abuse or eating disorders is very similar to the woman who does not wish the responsibility of constant hospital visits or supporting a

<sup>25</sup> T. Shakespeare, ‘Choices and Rights: eugenics, genetics and disability equality’, (1998) 13(5) *Disability & Society* 672

<sup>26</sup> E. Jackson, *Medical Law: Text, Cases And Materials* (OUP, Oxford 2006) 605

<sup>27</sup> Curt S Rush, ‘Genetic Screening, Eugenic Abortion and Roe v Wade: How Viable is Roe’s Viability Standard?’ (1983-1984) 50 *Brook L. Rev.* 113, 134

<sup>28</sup> *Abortion Time Limits: A Briefing Paper for the BMA* (BMA 2005)

<sup>29</sup> S. Vehmas, ‘Parental Responsibility and the Morality of Selective Abortion’ (2002) 5(4) *Ethical Theory and Moral Practice* 472

<sup>30</sup> *Ibid.*

child with limited mobility or cognitive ability for example. So once again we face the realisation that there is a stark difference in choice between contexts and it seems like an illogical distinction if one accepts, as Vehmas does, that “families of children with cognitive impairments do not necessarily experience any more difficulties than families with so-called normal children.”<sup>31</sup>

The logical answer, one might argue, actually lies in somehow restricting the scope of the social ground as opposed to a permissive attitude towards selective abortion. The rationale would be that if s.1(1)(a) were interpreted more stringently, if the concept of risk to the mental or physical health of the mother were somehow harder to satisfy then the disparity identified would be less significant. Consequently a ban on selective abortion would be more supportable. However, it is submitted that the social ground should be liberally interpreted in the interests of the mother. As Jackson points out, the Royal College of Obstetricians and Gynaecologists (RCOG) suggests that “health” in the Act refers to the World Health Organisation’s definition which includes “a state of physical and mental well-being.”<sup>32</sup> The well being of a woman who does not want to be pregnant would certainly be promoted by allowing her to have an abortion and an abortion would normally promote physical wellbeing as pregnancy and childbirth generally involve more risk than termination.<sup>33</sup>

One potentially devastating argument to the parental interest’s justification is that selective abortion could be discriminatory and that similar provisions for race and gender would not be accepted by society.<sup>34</sup> Essentially the argument is that surely it would be unacceptable that under legislation parents would be permitted to abort a child because it were black or female and might, due to societal conditions, be more difficult to raise/have a greater negative impact on the mother. There has been much debate about this issue but for the purposes of this article it will be dealt with swiftly by looking at the distinction between the social and impairment models of disability. The experiences of a blind person or someone with a cognitive disability are limited by his impairment notwithstanding the fact that some social conditions might make the disadvantage worse. This is in contrast to being female or black where there is no actual medical impairment, no intrinsic disadvantage. The only inability to do something that others can stems from social discrimination and prejudices.<sup>35</sup> Therefore the above argument does not negate autonomy and parental interests as a strong justification for selective abortion of disabled fetuses. Accepting the possibility of abortion for physical and mental impairments is not the equivalent of condoning the abortion of a female fetus purely on the grounds of its sex.

## **2. Legal Intervention and the Role of the Law: Current rationale, inconsistencies and proposed reform**

### ***2.1 Negligible Regulation of Selective Abortion – A step too far?***

Having concluded that none of the contemporary arguments against selective abortion justify banning it completely and that the autonomy and interests of the mother provide a positive reason to allow it, there remain important questions. What should be the role of

<sup>31</sup> Ibid

<sup>32</sup> Jackson (n26) 608

<sup>33</sup> Ibid

<sup>34</sup> Field (n14) 116. Sheldon and Wilkinson (n1) 100



the law and is it currently satisfactory? Consider the following view forwarded by Curt S Rush: “a state should not be permitted to prohibit eugenic abortions at any stage of a mother’s pregnancy.”<sup>36</sup> His only qualification is that regulation is necessary to protect the life or health of the mother. Rush argues that a severely defective child born must ordinarily endure considerable pain and suffering, referencing Tay Sachs.<sup>37</sup> If ‘severely’ is interpreted strictly then it could be said that the state has no right to regulate selective abortion based on the interests of the foetus even after 24 weeks. But surely it must be that Rush would need to accept that this justification could not prevent the regulation of Down’s syndrome or other conditions where it would be better for the child to be born at any point during the pregnancy? The reply to this question is that following the logic of Rush’s argument he need not accept such a position. This is so as he also asserts that the pregnant woman is the best person to judge how she will cope with the impact on her life of a severely disabled child and that the decision is ‘highly private’ and ‘so intimate and personal.’<sup>38</sup> Even if Down’s syndrome does not fall within his idea of ‘severely handicapped’ he states that “[n]o-one is better qualified”<sup>39</sup> to decide whether it would be right to allow the foetus into the world. The importance of personal autonomy and parental interests in justifying selective abortion has been stressed in this article. But it is submitted that reliance on this to assert unfettered control in selective abortion of severely disabled fetuses, as Rush does, is more problematic than helpful. Following this line of reasoning one must quickly concede that it calls for no regulation of abortions involving minor handicaps. It is arguable that the argument’s best justification is Mill’s harm principle – that persons should be free to live their lives without the interference of the state provided it does not harm anyone else.<sup>40</sup> So a mother, being the best judge of whether or not to have a child with, for example, cleft palate should be free to decide to have an abortion unless there is some actual harm caused to others. But recall the discussion in Part I. Selective abortion in general does not point to an existing class of disabled persons and say that it is better that they should not have been born at all nor does it marginalise/exclude these persons from society. The law refuses to accept that the foetus has the legal status of a person and Feinberg criticises the idea of ascribing to a foetus moral personhood. No actual persons are harmed and unless denying existence to future persons is a harm then it seems the mother should be free from interference. So the argument asserts that there is no problem with preventing regulation where the foetus is merely handicapped, as opposed to severely. However, the real problem with relying on Rush’s argument and the harm principle is that it does not stop there. The mother’s personal choice and discretion over how to live her life would surely always triumph in the normal abortion context. Surely the decision is always going to be ‘intimate and personal’ and ‘highly private’ regardless of whether the foetus is handicapped or not. And to whom is the harm directed? Whilst Rush would surely deny that this is what he was advocating it does seem to call into question the right to regulate normal abortions. Surely it challenges the right of the UK government to impose a 24-week limit by virtue of the Abortion Act 1967? Also, as is explored by Jackson, there is a question of how legitimate it is to require the approval of two doctors for all abortions, considering the mother is the best judge.<sup>41</sup> It would certainly seem to call for a

<sup>35</sup> Shakespeare (n25) 670. Sheldon and Wilkinson (n1) 104

<sup>36</sup> Rush (n27) 119

<sup>37</sup> Ibid, 134

<sup>38</sup> Ibid, 136

<sup>39</sup> Ibid

<sup>40</sup> JS Mill, *On Liberty* (Penguin Group, London 1974) 68

<sup>41</sup> E Jackson ‘Abortion, Autonomy & Prenatal Diagnosis’ (2000) 9(4) *Social & Legal Studies* 471.



right to abortion enshrined in the law. Essentially the problem is that whilst one might believe that autonomy legitimately bars the state from interfering in selective abortion involving severe handicap it is very difficult not to use this reasoning to bar state regulation of abortion under any circumstances.

## **2.2 The Right to Regulate?**

At this stage two main points will be argued for. It is submitted that 1) the state does have a legitimate interest in intervening in and regulating normal abortions; and 2) that this justification does not allow for selective abortion to escape regulation. In fact it will be demonstrated that based on the inherent right to regulate the current approach to selective abortion suffers from a dangerous incoherence. In essence it is inconsistent with the approach of the law to abortion in the non-disabled context.

What then are the state's justifications for interference? In the case of *Roe v Wade*, Justice Blackmun, delivering the opinion of the US Supreme Court, stated that a "pregnant woman cannot be isolated in her pregnancy" and that "a woman's privacy is no longer sole and any right of privacy she possesses must be measured accordingly."<sup>42</sup> It is proposed that this is the preferable view and is equally applicable to England and Wales. There must be some value placed on other considerations in addition to the woman's privacy and her autonomy. What are these considerations? At 28 and even 24 weeks it was said throughout the decision in *Roe* that the foetus has the capability of meaningful life outside the mother's womb. It is submitted that a similar rationale underlines the 24 week limit imposed by the Abortion Act 1967, the English law concept that the foetus is "capable of being born alive."<sup>43</sup> Blackmun J stated that the important and legitimate interest in protecting potentiality of life becomes "compelling."<sup>44</sup> This is linked to Feinberg's idea that the later on in the pregnancy the stronger the reason required to justify abortion of the foetus.<sup>45</sup>

## **2.3 Unnecessary Distinctions – The disabled vs. non-disabled foetus and the case of the neonate**

It is submitted, that if there is an ever growing reason not to abort as a normal, healthy foetus continually approaches personhood then it does not seem possible to deny it to foetuses that carry a substantial risk of serious handicap. It would surely be wrong to say that a foetus that has a defect has less moral worth than one that does not. It is the same save the handicap. Similarly the state has a general interest in preservation of life which grows stronger throughout the pregnancy notwithstanding the fact that the foetus is or is not handicapped. Of course as was also referred to in *Roe*, the state has an interest in the preservation and protection of the maternal health of the mother as well.<sup>46</sup> This is equally true of English law, encapsulated in ss.1(1)(b) and 1(1)(c) of the Abortion Act. These sections respectively allow for termination where it is necessary to prevent grave permanent injury to the physical/mental health of the mother and where the risk to her life is greater if the pregnancy is continued. Thus, it is important to remember that the foetus's personhood

<sup>42</sup> *Roe v Wade* (n10) 159

<sup>43</sup> Infant Life (Preservation) Act 1929 s 1 (1)

<sup>44</sup> *Roe v Wade* (n10) 163

<sup>45</sup> Feinberg (n11)

<sup>46</sup> *Roe v Wade* (n 10) 163

notwithstanding, its interest is not the only one at stake.

So, to frame it in the words of Feinberg, after 24 weeks is there a strong enough reason to allow selective abortion? Or to adopt the approach of the US Supreme Court in *Roe* can the compelling legitimate interest of the state in the preservation of life be outweighed? In reply to Feinberg it is stated that the autonomy of the mother and the parental interest's arguments are no longer strong enough reasons to allow selective abortion. Nor are these factors any longer capable of outweighing the legitimate interest the state has in the preservation of life espoused by the Supreme Court Justices in *Roe*. The mother's right to decide whether to continue with the pregnancy is the same in the case of a disabled or non-disabled foetus and so after 24 weeks this view should in each case have no more weight than in the other. Here one might argue that as discussed earlier, there may be increased strain and implications for the mental health of the mother. Indeed Rush points out that this adds to the reasons accompanying a normal pregnancy further aggravating the circumstances.<sup>47</sup> The British Medical Association was cited as support for the increased negative impact on the mother's health.<sup>48</sup> It was submitted in Part I that if this is necessarily the case then it provides a strong justification for allowing the practice of selective abortion. Nonetheless it does not follow that it justifies selective abortion at all points throughout the pregnancy. One might believe that this is unduly harsh on the mother. However, there is one consideration that could go a long way toward mitigating this perceived harshness. It is submitted that if continuing a pregnancy with a disabled foetus would be so detrimental to the health of the mother then there could still be a justification under the aforementioned s.1(1)(b), namely that it is "necessary to prevent grave permanent injury to the physical or mental health" of the woman. But one cannot lose sight of the fact that there is now a viable foetus, it is now capable of life outside the womb. As was argued above the moral worth of the foetus is the same in both the disabled and non-disabled contexts. As is the case with all non-disabled foetuses, the moral status of the handicapped foetus is paramount and it is accordingly now only going to be where it is in the best interests of the foetus to be aborted that selective abortion can take place. The state's interest in the preservation of life takes precedence. Abortion may still take place where there is a serious risk to the life of the mother but this is no longer selective abortion, the reason for the abortion is not the characteristics of the foetus but the preservation of the mother's life. Most people would find it uncontroversial that the mother must be allowed preferential treatment over the foetus, as disability is no longer a pertinent factor. But where disability is a pertinent factor and 24 weeks have passed the previously discredited foetal interest's argument is the only one capable of allowing selective abortion.

This position would consequently bring abortion law more into line with the treatment of neonates. Many cases have emphasised that only where the child's life would be so demonstrably awful<sup>49</sup> is the withdrawal of treatment justified. Only then is the severely handicapped infant allowed to die. So too would be the situation in the abortion context. The courts have asserted that the best interests test encompassing medical, emotional and all other welfare issues is preferable to some measure of intolerable or demonstrably awful life.<sup>50</sup> However, the idea of 'demonstrably awful' effectively conveys the gravity of the foetus's condition and the effect on the life of the future child. In recognition of this the

<sup>47</sup> Rush (n27)

<sup>48</sup> BMA 2005 (n 28)

<sup>49</sup> *Re B* [1981] 1 WLR 1421 CA 1424

<sup>50</sup> *Wyatt v Portsmouth Hospitals NHS Trust* [2005] EWCA Civ 4023

proposed legislation below incorporates both this term and reference to best interests. It is submitted that this relative parity between the treatment of neonates and fetuses after 24 weeks would be a welcome shift in the law. In *Re C*, a case concerning a very seriously disabled and terminally ill 16 month old child, the doctor's view was that "treatment would be futile, it would not improve her quality of life and would subject her to further suffering without... any benefit."<sup>51</sup> It becomes problematic to argue that the same considerations of welfare should not be taken into account as between a neonate and a post-24 week fetus that will develop into a child exhibiting these symptoms. This is so notwithstanding the fact that Feinberg would postulate the answer lies in personhood and that the fetus is neither a legal nor moral person.

Scott has suggested that as a:

Matter of ethics and law, that if it is in the best interests of a neonate to withdraw treatment it must also be justifiable to terminate the life of a fetus with a condition of the same degree of severity.<sup>52</sup>

This is surely correct. But one must clearly reject the assertion of Tom Shakespeare that to justify post-24 week selective abortion it must be "inevitable that the fetus will die before birth or in the neonatal period,"<sup>53</sup> as an unnecessary restriction. Selective abortion for conditions rendering life intolerable but which wouldn't lead to death during the neonatal period should surely be permissible. If the welfare of the fetus is the goal then it seems slightly counter-intuitive to say that it would only be where death was forthcoming that the fetus could be spared pain and suffering. For example where the potential infant would suffer from a condition rendering them blind, deaf and in pain but which would not be imminently fatal (indeed the child might live for many years) then an abortion should be available right up until birth.

## **2.4 Suggested Reform of Current Law**

The preceding discussion would translate into the following legislative amendments. Section 1(1)(d) as it currently stands would no longer represent the law regarding selective abortion. Firstly, where the mother wants to abort the child because it is, or there is a risk it will be, disabled then she can do so and there is no restriction based on substantial risk of a serious handicap. The logical question is by what legal mechanism would she do so? In answer to this it is submitted that she would seek the abortion under s.1(1)(a), asserting that there would be risk to her physical/mental health having consideration of the actual or reasonably foreseeable circumstances. Of course, amalgamating cases of selective abortion under that section would mean that the same restrictions apply. These are namely that the mother does not have complete control before 24 weeks as her decision is subject to two doctors' discretion and there is no legal *right* to abortion. The fact that there is a handicapped fetus would not change any of this. It is contended that this alteration to the law would be desirable as this article argues there is no real distinction before 24 weeks. Secondly, s.1(1)(d) would not be rendered irrelevant but it is proposed that it would undergo amendment and clarification to effectively serve as the justification for post-24

<sup>51</sup> [1998] Lloyd's Rep Med 1 Fam Div, 1

<sup>52</sup> Scott (n 4) 128

<sup>53</sup> Shakespeare (n17) 95

week selective abortion. The following provision, or one with similar wording, would be incorporated into the statute as the new s.1(1)(d):

That the pregnancy has exceeded its twenty-fourth week and that there is a substantial risk that if the child were born it would suffer from physical or mental abnormalities as to be “seriously handicapped”

In addition to this amendment a new subsection 2A would be inserted into the Act and would state the following:

“Seriously handicapped” in s.1(1)(d) means that the child’s life would be so demonstrably awful that it is in the best interests of the foetus for it to be aborted.

## **2.5 Consequences of Such Reform**

There are, however, several consequences that follow from the above legislative proposals and in conclusion consideration shall be given to a selection of them;

a) First of all, in contrast to the current situation “serious handicap” would now have a legal definition. The natural consequence would be that save for some highly unusual complicating factors in individual cases, the majority of conditions including cleft palate and Down’s syndrome would surely not fall under the definition of “serious handicap”. Therefore a woman would not be able to terminate such a foetus under the post-24 week disability ground. Whilst “serious handicap” has had no prior legal definition it is important to note that it has been interpreted by the RCOG who have produced guidelines to medical practitioners.<sup>54</sup> The impact of the amended Act for abortions after 24 weeks on these current guidelines should be examined for two very significant reasons. Firstly, despite the proposed legislative changes, a woman’s access to abortion will still be subject to the discretion of the medical profession acting in good faith. With no judicial interpretation the RCOG guidelines would still be the first port of call for doctors in interpreting the new act. Secondly, in considering the guidelines it could also (in addition to the neonate cases) prove a useful guide as to how the courts might also deal with the legal concepts of ‘demonstrably awful’ and ‘best interests’ in this context. Whilst the proposals of this Paper have substantially narrowed the concept of “serious handicap” there inevitably remains some ambiguity. However, these guidelines would be irrelevant to selectively aborting foetuses before 24 weeks as under the new regime there would not need to be a “serious handicap”.

One of the factors highlighted by the RCOG to take into account for determining whether a condition is a serious handicap is the probability of effective treatment *in utero* or after birth.<sup>55</sup> However, it would seem that in light of the proposed amendments, on the assumption that medical treatment meant remedying the affliction or even alleviating it considerably then this would be a decisive factor.<sup>56</sup> In other words where remedy or substantial alleviation was available then it would be highly problematic to establish that the condition could be serious. This is because the new s.1(1)(d) would only allow abortion

<sup>54</sup> RCOG, Termination of Pregnancy for Fetal Abnormality in England, Wales and Scotland (Jan 1996)

<sup>55</sup> Ibid 3.3.3

<sup>56</sup> Scott (n4) 80 for discussion of treatment/alleviation

where the life of the foetus would be “demonstrably awful” and this clearly won’t be the case where treatment is available and is unlikely to be so if substantial alleviation of the symptoms is possible.

Another factor is the suffering that would be experienced by the foetus.<sup>57</sup> It has been stated the RCOG guidelines are probably best understood as referring to physical suffering and not the result from some social stigma<sup>58</sup> and this, in conjunction with the neonate cases, would send the message to doctors (and courts) that the suffering would need to be tantamount to extreme, prolonged pain or uncomfortable living in order to satisfy the proposed law.

The RCOG also places emphasis on the probable degree of self-awareness.<sup>59</sup> However, it is quite conceivable that a low level of self-awareness would not amount to a life that were demonstrably awful and that it might still be in the best medical and emotional interests to opt for life over death. Of course if the situation were such that the foetus would only be aware of pain and could not appreciate any pleasure then this would almost undoubtedly satisfy the label of “serious”.

The two further RCOG considerations, “the extent to which the actions essential for health that normal individuals perform unaided would have to be provided by others” and “the probability of being able to live alone... be self-supporting as an adult”<sup>60</sup> can be taken together. Once again it is not inconsistent with a life worth living that much support and assistance would need to be administered to the child and a lifetime of care necessary. However, it may be that such extensive reliance on others for every basic need and no prospect of living alone combined with suffering and lack of awareness would strongly point away from carrying the foetus to term and qualify as ‘serious’. It is also worth highlighting that whilst under the current law these two guidelines do not preclude consideration by the medical profession of the effect on third parties in ascertaining seriousness<sup>61</sup> the proposals in this paper would. It is only where the life of the child would be demonstrably awful so that it would be in the best interests of the child that the requirement of “serious handicap” is met;

b) Whilst there has been much clarification on the meaning of ‘serious handicap’ under the suggested amendments there would still be confusion over exactly what amounts to a ‘substantial risk’ and this would need to be explored;

c) One particularly striking implication of the strict restriction of late abortions would be the impact on the decision making process of parents of the handicapped foetus. It must be conceded as inevitable that there may be increased stress in light of the Nuffield Report’s conclusion that some specialists believe the “absence of a cut-off point has relieved the pressure for a hurried decision.”<sup>62</sup>

The result will be in no small part determined by the professionalism of doctors and nurses in ensuring the decision making process is as informed and beneficial as possible since the parents will have less time to decide.

d) The proposed legislation creates a further significant problem. What about

<sup>57</sup> Ibid

<sup>58</sup> Scott (n4) 83

<sup>59</sup> RCOG, Jan 1996 (n54)

<sup>60</sup> Ibid

<sup>61</sup> Scott (n4) 85

<sup>62</sup> Nuffield Council on Bioethics, *Critical Care Decisions in Fetal and Neonatal Medicine: Ethical Issues* (2006) 4.13



abnormalities that can only be detected later in the pregnancy after the 24 week limit but could not conceivably fall within the new meaning of ‘*serious handicap*’ or where there wasn’t a ‘*substantial risk*’? Does this mean that the mother has lost the opportunity to abort? It could be seen as penalising the woman and forcing her to carry a child to term. One might argue she is essentially being made to suffer the consequences of the limitations of scientific medicine. This issue is problematic and there is no easy definite answer. Two considerations should be borne in mind. Firstly, while the mother has in essence lost an opportunity, there is now a viable foetus. It might still be right to say that the interests of the foetus cannot be overridden by anything less than the risk of serious harm or death to the mother. Secondly, Field questions whether allowing selective abortion in the late diagnosis situation might lead to allowing late abortions in general based on reasons such as a woman’s husband dying late in pregnancy or where her boyfriend leaves her.<sup>63</sup> The justification for such an idea in English law would seem to be under s.1(1)(a), that different circumstances now pose a risk to her mental health greater than carrying on the pregnancy. When expressed in those terms it is not obvious why theoretically this would not be allowed if an exception were made for conditions detected later.

e) It would also follow that *prima facie* any changes to the 24 week limit for the social ground should carry over into the disability ground. In other words since my argument is based on the need for uniformity when it comes to disabled and non-disabled foetuses (and parental interests) any decrease in the 24 week limit for the social ground should be matched by an appropriate reduction in the time that needs to have passed for the new s.1(1)(d) to kick in. It is a rebuttable presumption in light of a definitive conclusion on the aforementioned detectability issue. If some exception were made for conditions detected later then obviously a reduction from 24 weeks wouldn’t effect all later terminations.

## Conclusion

In closing, it has been argued that it would be wholly inappropriate to ban selective abortion and that to permit it *does not* offend the dignity and integrity of the foetus nor the disabled community whilst it *does* respect the mother’s right to self-determination. It is desirable for the state to be involved and in light of this legislation, albeit far from perfect, has been proposed to eliminate current inconsistencies in what are, it is submitted, like situations.

<sup>63</sup> Field (n14) 114





# PROCREATIVE LIBERTY AND SELECTING FOR DISABILITY: SECTION 14(4) HUMAN FERTILISATION AND EMBRYOLOGY ACT 2008

EDWARD M TAYLOR, LLB  
KING'S COLLEGE LONDON  
EDWARDMTAYLOR@GMAIL.COM

**ABSTRACT** On October 1st 2009, s.14(4) of the Human Fertilisation and Embryology Act 2008 amended the Human Fertilisation and Embryology Act 1990 to expressly prohibit the selection of disabled embryos for implantation. The first part of this article examines how s.14(4) has altered the law, and in turn the regulatory system, to prevent parents from using reproductive technologies, such as Preimplantation Genetic Diagnosis, to predetermine the birth of a disabled child (selecting for disability). Section 14(4) is seen to have placed new restrictions upon the Human Fertilisation and Embryology Authority's (HFEA) previously discretionary licensing powers. The second part of this article utilises the principle of Procreative Liberty to consider whether s.14(4) can be justified: does selecting for disability harm the resulting disabled child? It is argued that Derek Parfit's 'non-identity problem' prevents the seemingly sensible claim that harm is caused in a person-affecting sense, except in relation to the small number of disabilities that are so severe as to be not compatible with a life worth living. Non person-affecting conceptions of harm are also found to be incapable of justifying s.14(4). On the basis of this analysis, the third part of this article argues that a total prohibition on selecting for disability, such as s.14(4), cannot be justified and criticises the government for failing to provide a coherent justification for s.14(4). A flexible regulatory approach based on the reinstatement of the HFEA's discretionary licensing powers is suggested as the most appropriate method of handling the difficult issues raised by selecting for disability.

## 1. Introduction

In 2002, Sharon Duchesneau and Candy McCullough's efforts to have a deaf child in the United States garnered widespread, largely critical, international media coverage.<sup>1</sup> The couple's story<sup>2</sup> revealed that a small number of disabled parents might want to use reproductive technologies, such as Preimplantation Genetic Diagnosis (PGD), to ensure the birth of disabled children (selecting for disability). PGD is a "technique to determine genetic defects in embryos created by in vitro fertilisation (IVF) prior to implantation in a uterus for gestation."<sup>3</sup> The technique is normally used to screen-out genetic defects, by selecting unaffected embryos<sup>4</sup> to implant into the woman's uterus, ensuring the birth of non-disabled children. Conversely, PGD can also be used to select for disability, through the selection and implantation of affected embryos, ensuring the birth of disabled children. A recent study has shown that "5% of 190 PGD clinics surveyed in the US have allowed parents to select embryos with conditions commonly taken to be disabilities."<sup>5</sup>

Section 14(4) of the Human Fertilisation and Embryology Act 2008 (HFE Act 2008) expressly prohibits clinics from using PGD to select for disability in the United Kingdom.<sup>6</sup> This article utilises the principle of Procreative Liberty, which states that people should have the "fundamental right"<sup>7</sup> to "decide whether or not to have offspring and to control the use of one's reproductive capacity,"<sup>8</sup> to consider whether the s.14(4) prohibition on selecting for disability can be justified. Part 1 outlines a scenario where a couple want to select for disability (the Condition X Scenario). This scenario is used in Part 3 to consider how Section 14(4) has changed the law in relation to selecting for disability. Part 4 explains the significance ascribed to Procreative Liberty in the regulation of reproductive technologies, and Parts 5 and 6 use its tenets to consider whether the s.14(4) prohibition can be justified. Finally, Part 7 considers an alternative approach to the regulation of selecting for disability.

## 2. PGD and Selecting for Disability: the Condition X Scenario

To provide a starting point for this article, it is useful to outline a hypothetical scenario (the Condition X Scenario) in which a couple want to select for disability: consider where the couple both have Condition X (a notional disability) and want to ensure that their child also has the condition. IVF is used to create 4 embryos, which are tested using PGD. The PGD reveals two embryos with Condition X and two unaffected embryos. The couple want to implant an embryo with Condition X. The legality of selecting for disability in the Condition X Scenario will be considered in Part 3. For now, it is necessary to outline three premises upon which this article proceeds. The first is that this article does not accept the

<sup>1</sup> The deaf lesbian couple used a deaf sperm donor, rather than PGD; their baby, Gauvin, was born partially deaf. The Washington Post Magazine (March 31st, 2002); BBC (8th April, 2002) <http://news.bbc.co.uk/1/hi/health/1916462.stm>.

<sup>2</sup> "Couple" is the generic name used for PGD applicants.

<sup>3</sup> Bonnie Steinbock 'Preimplantation genetic diagnosis and embryo selection' (2002), in *A Companion to Genetics* eds J. Harris and J. Burley (Oxford, 2002) p.175.

<sup>4</sup> "Unaffected embryos" refers to embryos PGD has shown not to have the screened-for disability.

<sup>5</sup> Julian Savulescu & Guy Kahane 'The Moral Obligation to Create Children with the Best Chance of the Best Life' *Bioethics* 23 (5): 274-290 2009 p.284 citing a survey in S. Baruch, D. Kaufman & K.L. Hudson 'Genetic testing of embryos: practices and perspectives of US IVF Clinics' *Fertil Steril* 2008; 89: 1053-1058 note 6.

<sup>6</sup> HFEA 2008: Explanatory Notes; paragraph 114. Section 14(4) also applies to gamete donation.

<sup>7</sup> John Harris 'Sex selection and regulated hatred' *Journal of Medical Ethics* 2005 vol. 31: p 287.

<sup>8</sup> John A. Robertson, *Children of Choice* (Princeton, 1994) p 16.

force of the general moral arguments against PGD use, which include the alleged immorality of selection and the supposed dangers of interfering with nature, and will not consider them further because the author considers that they fail to show that PGD should be subject to a general prohibition.<sup>9</sup> Indeed, this article is only concerned with moral arguments to the extent that they assist in determining whether the s.14(4) prohibition on selecting for disability breaches the principle of Procreative Liberty.

The second premise concerns the relationship between disability and quality of life.<sup>10</sup> It is clear that disability does not correspond directly to quality of life; the same disability will affect different people in different ways.<sup>11</sup> Indeed, big ‘D’ deaf rights activists assert that deafness is a “linguistic and cultural identity,”<sup>12</sup> rather than a disability, which improves their quality of life.<sup>13</sup> Nevertheless, this article proceeds on the basis that there is a “predictable and straightforward relationship between disability and quality of life, such that being disabled is a reliable predictor of lower quality of life.”<sup>14</sup> Thus, selecting for disability will be deemed to create children with decreased life prospects.

The third premise is that Procreative Liberty provides a useful benchmark against which the legitimacy of s.14(4) can be measured. The prominence given to Procreative Liberty by the bodies charged with regulating reproductive technologies, as will be seen in Part 4.1, shows that this is not an unreasonable premise. It is to be noted, however, that the broader debate as to whether Procreative Liberty should be used, in isolation, as the benchmark against which reproductive technologies should be measured is beyond the scope of this article.<sup>15</sup>

### 3. The Law: Selecting for Disability

The Condition X Scenario will be used throughout this Part to consider how the HFE Act 2008 has affected the HFEA's ability to license selecting for disability. Part 3.1 provides an overview of the regulatory framework in relation to PGD, before Parts 3.2 and 3.3 consider respectively the ‘old law’ of the unamended Human Fertilisation and Embryology Act 1990 (HFE Act 1990) and the ‘new law’ of s.14(4) HFE Act 2008.

#### 3.1 Regulating PGD: the HFEA, HFE Act 1990 and HFE Act 2008

PGD and other reproductive technologies are regulated by the HFEA; an independent statutory authority, responsible for licensing and monitoring UK clinics offering fertility services.<sup>16</sup> Clinics can only offer treatments that the HFEA has licensed. The HFEA's statutory framework was created by the HFE Act 1990. Prior to the HFE Act 2008, the HFEA had a significant discretion when deciding whether to license PGD for new uses, subject to the general licensing conditions of ss.12-15 HFE Act 1990. This discretion stemmed from the fact that the HFE Act 1990 did not impose express statutory prohibitions

<sup>9</sup> For a similar conclusion see: John A. Robertson ‘Ethics and the Future of Preimplantation Genetic Diagnosis’, *Reproductive BioMedicine Online* Vol 10 (2005) p.98-99; Holm S. “Ethical Issues in Pre-implantation Diagnosis” p.190 contained in *The Future of Human Reproduction: Ethics, Choice, and Regulation* (Oxford, 1998).

<sup>10</sup> Stephen Wilkinson with Eve Garrard ‘Selecting Disability and Welfare of the Child’ (2006) *The Monist*, 89 p 1.

<sup>11</sup> *Ibid.* p 2.

<sup>12</sup> Dena S Davis. *Genetic Dilemmas: Reproductive Technology, Parental Choices and Children's Future* (Routledge, 2001) p 52.

<sup>13</sup> Wilkinson (n 10) p 1 footnote 414 Wilkinson (n 10) p 2.

<sup>14</sup> Wilkinson (n 10) p 2.

<sup>15</sup> For an interesting argument in favour of setting a low benchmark for the regulation of reproductive technologies see: Colin Gavaghan, *Defending the Genetic Supermarket* (Routledge-Cavendish, 2007).

<sup>16</sup> Review of the Human Fertilisation and Embryology Act: A Public Consultation, Department of Health (2005); para 2.10 (p 12).

on PGD. The HFEA nevertheless decided to adopt a relatively restrictive approach to the regulation of PGD; clinics were only licensed to ‘screen-out’ embryos with a “significant risk of serious genetic conditions.”<sup>17</sup>

The HFE Act 2008 amended the HFE Act 1990 to place express statutory prohibitions, such as s.14(4), on the use of PGD. These amendments have significantly reduced the HFEA’s discretion, but will not have a significant impact on the availability of PGD at the point of use as the amendments have been largely limited to placing the HFEA’s restrictive approach to regulating PGD on a statutory footing.

### **3.2 HFE Act 1990 (unamended): Discretionary Powers and Section 13(5)**

To receive PGD, the couple in the Condition X Scenario would apply to a fertility clinic licensed by the HFEA.<sup>18</sup> The clinic would likely refer the matter directly to the HFEA Licence Committee; a procedure followed where clinics receive applications for other controversial uses of PGD such as HLA tissue typing and late onset conditions.<sup>19</sup>

Under the ‘old law’, the HFEA Licensing Committee would consider “the unique circumstances of those seeking treatment”<sup>20</sup> in deciding whether to license PGD for the Condition X Scenario. The HFEA’s general approach has been to only license clinics to screen-out embryos with a significant risk of serious genetic conditions.<sup>21</sup> Indeed, the HFEA has not licensed PGD for ‘screening-in’ any traits,<sup>22</sup> and “[c]urrent HFEA licenses for PGD typically forbid the implantation of ‘abnormal’ embryos.”<sup>23</sup> Thus, the HFEA would be unlikely to grant a licence to select for disability.

Nevertheless, the HFEA could potentially have exercised its discretion to licence the Condition X Scenario. The HFEA’s discretion would only be limited if selecting for disability breached a general licensing requirement of ss.12-15.

On applying the general licensing conditions, s.13(5) could potentially be engaged: “A woman shall not be provided with treatment services unless account has been taken of the welfare of any child who may be born as a result of the treatment...”<sup>24</sup> Thus, the HFEA would need to consider whether selecting for disability is detrimental to the welfare of the child. The Code of Practice guidance for s.13(5) outlines ‘relevant risk factors,’<sup>25</sup> which include “... any aspect of the patient’s (or, where applicable, their partner’s) medical history which means that the child to be born is likely to suffer from a serious medical condition.”<sup>26</sup> If Condition X qualified as a serious medical condition, the Code of Practice’s interpretation of s.13(5) suggests that selecting for Condition X would be deemed detrimental to the welfare of the child.

<sup>17</sup> HFEA Code of Practice, 7th Edition (G.12.3.2). The HFEA’s Code of Practice “gives guidance to fertility clinics about the proper conduct of HFEA-licensed activities”. ‘About the Code’, (accessed 2nd March: [www.hfea.gov.uk/en/371.html](http://www.hfea.gov.uk/en/371.html)).

<sup>18</sup> A clinic has to have a general HFEA license and a license for individual genetic conditions (HFEA 1990 s.11 and Schedule 2). In the absence of a licence, an offence is committed (HFEA 1990 s.41(1) and (2)).

<sup>19</sup> HFEA <http://www.hfea.gov.uk/docs/PreimplantationReport.pdf>.

<sup>20</sup> Code 7th (n17) G.12.3.1.

<sup>21</sup> Code 7th (n17) G.12.3.3.

<sup>22</sup> Code 7th (n17) G.12.5.1- G.12.5.8. Tissue typing is allowed in limited circumstances, and might be considered ‘screening in.’

<sup>23</sup> Rosamund Scot, *Choosing Between Possible Lives Law and Ethics of Prenatal and Preimplantation Genetic Diagnosis* (Oxford, 2007) p 312.

<sup>24</sup> For s.13(5) critique: Emily Jackson, “Conception and the Irrelevance of the Welfare Principle.” *Modern Law Review* (2002) Vol. 65, p 176-203.

<sup>25</sup> Code 7th (n17) G.3.3.2.

<sup>26</sup> Code 7th (n17) G.3.3.2(b).

On this basis, s.13(5) would prevent the HFEA from granting licenses to select for disability. The claim that selecting for disability harms the child underpins the majority of arguments favouring a s.14(4) type prohibition on selecting for disability. It will be seen in Part 6.1, however, that selecting for disability cannot be said to harm the child, or be detrimental to its welfare except in the case of the most serious disabilities.<sup>27</sup> Thus, s.13(5) should not place a general prohibition on the HFEA licensing selecting for disability cases.

To conclude, the HFEA's restrictive approach to licensing PGD would have prevented the grant of a licence to select for disability in the Condition X Scenario. They would have had the discretion to do so, however, as s.13(5) sensibly does not apply in all such cases.

### 3.3 HFE Act 2008: Section 14(4)

Section 14(4) of the HFE Act 2008 inserts ss.13(8) and 13(9) into s.13 HFE Act 1990: an embryo "known" to have a "serious physical or mental disability... must not be preferred to those that are not known to have such an abnormality," when deciding "which of two or more embryos to place in a woman." This amendment to the general licensing conditions means that embryos with serious conditions cannot be preferred over embryos not known to have such conditions.

Section 14(4) appears to leave open the possibility of selecting for non-serious conditions. This is prohibited, however, by other HFE Act 2008 amendments. In particular, the amended Schedule 2 1ZA(2) states that embryos can only be tested for PGD where there is the potential for a "serious physical or mental disability, a serious illness or any other serious medical condition."<sup>28</sup> Indeed, the Government's rationale for the HFE Act 2008 was partly to address "concerns that... PGD could be used to avoid less serious conditions."<sup>29</sup> Thus, PGD can only be initially authorized if the condition is "serious", preventing the selection of non-serious conditions.

Returning to s.14(4), the HFEA's proposed interpretation of 'preferred' does allow some flexibility regarding the implantation of disabled embryos.<sup>30</sup> If a couple sought to screen out Condition X, but PGD revealed that all the viable embryos had the condition, s.14(4) would not prevent the implantation of the disabled embryos as they are not being 'preferred.' This is a situation "where the only unaffected embryos are of such low quality that they are unlikely to result in a pregnancy."<sup>31</sup> This interpretation leaves room for "clinical decision making."<sup>32</sup>

As an aside, s.14(4) appears to have clarified the law on a difficult scenario: where the couple suffer from Condition X, but only want to use PGD to screen their embryos for Condition Y. On applying the 'old law' to a similar situation, Scott concludes the Code of Practice might oblige clinical staff to also test the embryos for Condition X,<sup>33</sup> despite the couple's wishes.

<sup>27</sup> Derek Parfit, *Reasons and Person* (Oxford, 1986) Ch. 16; Dan W. Brock 'The Non-Identity Problem and Genetic Harms – The Case of Wrongful Handicaps' *Bioethics* Vol. 9 (1995).

<sup>28</sup> HFE Act 2008: Schedule 2 Clause 3.

<sup>29</sup> HFEA Regulating preimplantation genetic diagnosis for the future: Listening to your views (28th January 2009) (2009) p 5.

<sup>30</sup> HFEA Code of Practice 8th Edition p.76 10C: 'Where there is no other embryo suitable for transfer, an embryo with these characteristics may be transferred'; HFEA Listening (n29) p 8-9.

<sup>31</sup> HFEA Listening (n29) p 8.

<sup>32</sup> HFEA Listening (n29) p 8.

<sup>33</sup> Scott (n23) p 313.



At a recent meeting, Tedd Webb, the Deputy Director of the Department of Health, suggested that s.14(4) would not lead to such a result. He stated that “PGD would be for condition Y, and therefore condition X wouldn’t come into it. The PGD would spot condition Y, not condition X.”<sup>34</sup> As the clinician does not ‘know’ whether the embryo has condition X, s.14(4) does not prevent him from inserting it, nor oblige him to test for it. This approach is sensible; requiring a clinician to test for Condition X would inevitably raise concerns of eugenics and discrimination.

To conclude, s.14(4) and other HFE Act 2008 amendments have prohibited PGD from being used to ‘select for disability.’ Thus, the couple would not be able to select an embryo with Condition X, unless it represented the only viable embryo. The HFE Act 2008 has significantly reduced the HFEA’s discretion to licence PGD. Part 4 utilises Procreative Liberty to consider whether the s.14(4) prohibition on selecting for disability can be justified.

## 4. Regulating Selecting for Disability: Procreative Liberty and Section 14(4)

### 4.1 Procreative Liberty and the regulation of reproductive technologies

Procreative Liberty provides that people should have the “fundamental right”<sup>35</sup> to “decide whether or not to have offspring and to control the use of ones reproductive capacity.”<sup>36</sup> This freedom is particularly significant in the field of procreation because of “the great importance to individuals of having biologic offspring – personal meaning in one’s life, connection with the future generations and the pleasures of child rearing.”<sup>37</sup> As reproductive decision-making frequently “turns on the expected child-rearing experiences that reproduction will bring,” Procreative Liberty allows couples “to obtain and act on information about a prospective child’s health and make-up in deciding whether or not to reproduce.”<sup>38</sup>

Thus, Procreative Liberty provides citizens with a *prima facie* right to access reproductive technology such as PGD; a right enjoying ‘presumptive priority’<sup>39</sup> over other concerns that might seek to proscribe this freedom.<sup>40</sup> The importance of Procreative Liberty in the regulation of reproductive technologies has been recognised by the HFEA,<sup>41</sup> the Human Genetics Commission,<sup>42</sup> the House of Commons Science and Technology Committee,<sup>43</sup> and academics.<sup>44</sup>

Procreative Liberty is not an unlimited right. It does not apply in situations where the use of the reproductive technology would cause “substantial harm to the tangible interests of others.”<sup>45</sup> It must also be noted that Procreative Liberty only acts as a legal principle, to

<sup>34</sup> Progress Educational Trust: ‘Debating Deafness and Embryo Selection: Are We Undermining Reproductive Confidence in the Deaf Community?’

<sup>35</sup> Harris (n 7).

<sup>36</sup> Robertson (n 8).

<sup>37</sup> Robertson (n 8) p 152.

<sup>38</sup> Robertson (n 9) p 99.

<sup>39</sup> Robertson (n 8) p 16.

<sup>40</sup> Wolfendon Report (1957), influenced by HLA Hart *Law, Liberty and Morality*; United States Constitution 14th Amendment.

<sup>41</sup> HFEA ‘Sex selection: options for regulation. A report on the HFEA’s 2002-03 review of sex selection’ (2003) para 132.

<sup>42</sup> Human Genetics Commission: *Making Babies* (January 2006) [1.5].

<sup>43</sup> House of Commons Science and Technology Committee: *Human Reproductive Technologies and the Law*, Fifth Report of Session 2004-05, Volume 1 (2005) [144-145].

<sup>44</sup> Robertson (n 8); Harris (n 7); Julian Savulescu ‘Deaf Lesbians, designer disability and the future of medicine’ *BMJ* (2002) 771-773; Ronald Dworkin, *Life’s Dominion* (Harper Collins, 1993).

<sup>45</sup> Robertson (n 8) p 41; John Stuart Mill, *On Liberty* (1859) p 21-22 ‘Harm Principle’.

guide the regulation of reproductive technology, not as a moral principle.<sup>46</sup> Thus, one can be morally opposed to selecting for disability, whilst recognising that citizens should have the right to make this choice when it does not cause substantial harm.

Finally, Procreative Liberty is only a “negative” right. The state, and others, have a duty not to interfere with a citizen’s right to exercise these choices, but are not obliged to provide the resources to allow citizens to exercise that choice.<sup>47</sup> Indeed, Procreative Liberty covers “little ground,” only revealing the freedoms that “cannot be criminalised.”<sup>48</sup> It nevertheless provides a useful benchmark against which the legitimacy of s.14(4) can be measured.

## 4.2 Procreative Liberty and Section 14(4)

The task of applying Procreative Liberty to selecting for disability is complicated by the fact that there is “no widespread agreement as to the nature and scope of this right.”<sup>49</sup> Fortunately, an adapted version of Robertson’s decisional methodology provides a useful test for deciding whether Procreative Liberty applies to selecting for disability. Under this test, s.14(4) will be justified if: (i) selecting for disability falls outside the scope of Procreative Liberty (whatever its scope and nature may be); and/or (ii) selecting for disability causes “substantial harm.”<sup>50</sup> These questions are considered in turn.

### Internal Constraints: “Is selecting for disability the type of decision that falls within the scope of Procreative Liberty?”<sup>51</sup>

The task of determining whether selecting for disability falls within the scope of Procreative Liberty depends, according to John Robertson, “on an evaluation of the importance of the choice to the parents and whether that choice plausibly falls within societal understandings of parental needs and choice in reproducing and raising children.”<sup>52</sup> This two-limb test, comprising a subjective and objective standard respectively, acts as an “internal constraint” on Procreative Liberty; those choices falling outside its scope are not protected.

The first limb requires a subjective evaluation of whether selecting for disability is important to the disabled couple’s reproductive decision making.<sup>53</sup> It will be suggested that selecting for disability could be important in four ways: emotional, cultural, practical and medical. This is not intended to be an exhaustive list of considerations, but provides a useful starting point for an enquiry.

To explore these factors, a hypothetical scenario involving an achondroplastic couple will be used.<sup>54</sup> The couple view achondroplasia as an important part of their lives. They feel uncomfortable with the idea of raising a non-achondroplastic child, who would not share their daily experiences. This leads them to believe that a stronger emotional relationship would be had with an achondroplastic child. Moreover, the couple view achondroplasia as a

<sup>46</sup> Shelley Kagan, *The Limits of Morality* (Oxford, 1989).

<sup>47</sup> Robertson (n 8) p 23.

<sup>48</sup> Davis (n 12) p 41.

<sup>49</sup> Harris (n 7) p 293.

<sup>50</sup> Robertson (n 9) p 99.

<sup>51</sup> Robertson (n 9) p 99.

<sup>52</sup> John A. Robertson, ‘Extending preimplantation genetic diagnosis: medical and non-medical uses.’ *Journal of Medical Ethics* (2003) 29(4) p 216, cited in Gavaghan (n 15) p 19.

<sup>53</sup> A disabled couple is used because it seems unlikely that non-disabled couples would want to use PGD to select for disability. Indeed, such selection could only occur if a disabled embryo was present.

<sup>54</sup> For detailed discussion: Davis (n 13) p 1-2, 49-51.

culture, rather than a disability, that they would like to share with their child.<sup>55</sup>

The couple's desire to have an achondroplastic child is also linked to practical concerns, as the potential size disparity of a non-achondroplastic child leads the couple to worry about maintaining discipline and how the child would cope with their adapted house. On medical grounds, the couple know that a non-achondroplastic pregnancy could have serious health implications for the woman.<sup>56</sup> They are also aware that non-assisted reproduction would carry a one in four chance of the resulting child having a "double-dose" of the achondroplastic gene,<sup>57</sup> which would cause sudden death. The couple don't want to take these risks.

These factors show that the ability to select an achondroplastic child is of paramount importance to the couple's reproductive decision making, as it will have a significant impact on their child-rearing experience. Indeed, the medical risks are such that denying the couple a right to select for achondroplasia would effectively deny them a child. This scenario does not show that selecting for disability will always be important to parents who want to select for disability; this will depend on the facts of each case. Rather, it shows that selecting for disability could be important, and thus satisfy the first limb of the test.

The second limb of Robertson's test requires an objective evaluation of whether selecting for disability "plausibly falls within societal understanding of parental needs and choice in reproducing and raising children." This raises the difficult question of where to draw the boundaries of societal understanding.<sup>58</sup>

The HFEA's approach has been to set societal understanding at a high level;<sup>59</sup> Procreative Liberty will only protect the right to make choices that the majority of society supports. On this approach, Procreative Liberty would protect the right to use PGD to screen out disabilities, but would not recognise "unpopular" choices such as "intentional diminishment of offspring traits [selecting for disability],"<sup>60</sup> or even sex selection.<sup>61</sup> If this standard were to be employed, s.14(4) would represent a legitimate prohibition.

Such an approach can be heavily criticised, however, for undermining the very freedom that Procreative Liberty is intended to protect. The freedom to make a choice only so far as it is accepted by society does not seem to be freedom at all.<sup>62</sup> This criticism can be partially avoided if "societal understanding" is instead set at a low level. Indeed, Robertson appears to advocate such an approach when he recognises that "the case for condemning PGD use to select for deafness is weak."<sup>63</sup> This standard would cast doubt on the legitimacy of s.14(4), as some instances of selecting for disability, such as deafness or achondroplasia, would seem to fall within Procreative Liberty.

Although this "weak" objective standard is defensible, it represents a difficult line-drawing exercise, prone to criticisms for arbitrariness. Moreover, it is unclear why society should have the right to limit an individual's "fundamental right" beyond the substantial harm limitation. For these reasons, this article takes the view that the second objective limb

<sup>55</sup> Davis (n 12) p 52.

<sup>56</sup> Scott (n 23) p 50.

<sup>57</sup> Davis (n 12) p 2.

<sup>58</sup> For a detailed critique re. objective nature: Gavaghan (n 16) p.21.

<sup>59</sup> HFEA 'Sex Selection' (n 41); Harris (n 7) p.292-294 cf. Tom Baldwin "Reproductive liberty and elitist contempt: reply to John Harris" *Journal of Medical Ethics* 2005 Vol.31 p 288-290.

<sup>60</sup> Robertson (n 9) p 99.

<sup>61</sup> HFEA 'Sex selection' (n 41).

<sup>62</sup> Gavaghan (n 15) p 20.

<sup>63</sup> Robertson (n 9) p 101.

of Robertson's test is not appropriate: "there is a limit to the legitimate interference of collective opinion with individual independence."<sup>64</sup>

To conclude, the "internal constraint" on Procreative Liberty requires selecting for disability to be subjectively important to the parent's reproductive decision making. Thus, s.14(4) will only be legitimate if all cases of selecting for disability are shown to cause substantial harm.<sup>65</sup>

**Substantial harm: "If so, would selecting for disability impose substantial harm on others that justify discouraging or barring the exercise of that choice?"<sup>66</sup>**

Procreative Liberty will not protect the right to exercise a choice causing "substantial harm to others." It is clear that "harm" in this sense does not include "harms to personal conceptions of morality, right order, or offence."<sup>67</sup> Rather, harm must be caused "to the tangible interests of others,"<sup>68</sup> which includes both physical harm and harm to resources.<sup>69</sup> As Gavaghan notes, there is a "world of difference between decisions of which we disapprove, and decisions which we would, or should, feel justified in prohibiting."<sup>70</sup>

It is important to note that Procreative Liberty's "presumptive priority" places the burden of proving the existence of substantial harm on the Government; "the onus should be on those who oppose... using PGD to show harm from its use."<sup>71</sup> Part 5 considers whether the Government has been able to justify s.14(4) by proving that selecting for disability causes physical harm to persons and/or harm to resources.

## 5. The Government: Proof of substantial harm?

In December 2006, the Government's White Paper on the reform of the HFE Act 1990 stated that "deliberately screening-in a disease or disorder will be prohibited."<sup>72</sup> This restriction was subsequently incorporated into Clause 14(4) of the Human Tissue and Embryos (Draft) Bill.<sup>73</sup> The Government failed to justify s.14(4) during the HFE Act 1990 review process and it is therefore necessary to consider any justifications provided during the parliamentary process.

In the House of Commons, the Secretary of State for Health, Alan Johnson, recognised that Clause 14(4) would not "permit couples to screen embryos in order to include, rather than exclude, a particular disability,"<sup>74</sup> but did not explain why this was a justifiable restriction.

Earl Howe was alone in clearly justifying his support for s.14(4): "I find that idea [of selecting for disability] repellent because it ignores one of the issues central to any IVF procedure, namely, the future welfare of the child."<sup>75</sup> This view suggests that selecting for

<sup>64</sup> Mill (n 45) p 130, cited by Gavaghan (n 16) p 36.

<sup>65</sup> The wider implications of such an approach to the HFEA's regulation of reproductive technologies is beyond the scope of this Article, but are briefly referred to in the conclusion.

<sup>66</sup> Robertson (n 9) p 99.

<sup>67</sup> Robertson (n 8) p 46.

<sup>68</sup> Robertson (n 8) p 41.

<sup>69</sup> Cf radical liberal; only accept physical harm to persons.

<sup>70</sup> Gavaghan (n 15) p 4.

<sup>71</sup> House of Commons Science and Technology (n 44) [144-145].

<sup>72</sup> Government's White Paper on Proposals for Revised Legislation (December 2006), paragraph 2.43.

<sup>73</sup> *Human Tissue and Embryos (Draft) Bill* (May 2007); prohibition was accepted by Joint Committee on the Human Tissue and Embryos (Draft) Bill: Volume I: Report (1 August 2007) p 54 (para. 195); p 88 "... nor is it permitted to select specifically for an abnormality, such as deliberately choosing an embryo which would result in a deaf child."

<sup>74</sup> Commons Hansard: 12 May 2008 1070.

<sup>75</sup> Lords Hansard 21 January 2008 23.

<sup>76</sup> Commons Hansard 19th May 2008 23.

disability physically harms the child. Mr. Leigh also voiced support for the prohibition, but on the rather vague ground that selecting for disability is “ethically dangerous.”<sup>76</sup>

In contrast, several Deaf organizations, such as the British Deaf Association, the World Federation of the Deaf<sup>77</sup> and the [stopeugenics.org](http://stopeugenics.org) website,<sup>78</sup> publicly criticised s.14(4) for restricting “reproductive liberty for people with specific characteristics and preventing the birth of certain kinds of people, including deaf people.”<sup>79</sup> These criticisms were noted at parliamentary questions<sup>80</sup> and the Department of Health went so far as to meet with the groups,<sup>81</sup> but this did not lead to parliamentary debate on the issue.

The lack of debate is particularly surprising given that the HFE Act 1990 review process did not reach a unanimous conclusion on the issue of selecting for disability. The House of Commons Science and Technology Committee<sup>82</sup> and some respondents to the Department of Health’s public consultation,<sup>83</sup> had in fact suggested that selecting for disability should not be prohibited (see Part 7.1, *infra*).

The Government’s grounds for s.14(4) would be unknown if it were not for their response to an online petition<sup>84</sup> opposing the prohibition: “it is in the best interests of the child not to prefer embryos that have a significant risk of developing a serious medical condition”.<sup>85</sup> Thus, it would seem the Government’s justification for s.14(4) rests on the idea that selecting for disability physically harms the child; a view that has been seen to be shared by Earl Howe and the HFEA in relation to s.13(5) (see Part 3.2, *supra*).

To conclude, the Government appears to have sought to justify s.14(4) on the premise that selecting for disability is detrimental to “best interests” or “welfare”, which can be seen as analogous to claiming that it causes “substantial” physical harm to the disabled child. The legitimacy of this claim, and possible arguments relating to harm to resources, are considered in Part 6.

## 6. Substantial Harm? Physical Harm and Resources

In Part 4.2, “substantial harm” was seen to consist of both physical harm and harm to resources. The Government’s justification for s.14(4) rests solely on the claim that selecting for disability causes physical harm.<sup>86</sup> This is perhaps unsurprising as the most “popular and intuitive answers to the question of ‘what’s wrong with selecting for disability?’ appeal to the welfare of the child.”<sup>87</sup> Nevertheless, this section also considers the “highly sensitive” issue of whether selecting for disability can be said to harm resources.<sup>88</sup>

<sup>77</sup> <http://stopeugenics.org/files/2008/04/ukbill-letter1.pdf>.

<sup>78</sup> <http://stopeugenics.org>.

<sup>79</sup> For summary of s.14(4) concerns: Edward White ‘Human Fertilisation and Embryology Bill [HL] Bill 70 of 80 2007-08’ Research paper 08/42 (2nd May 2008): p.21-23.

<sup>81</sup> Commons Hansard April 2nd, 2008 1094W.

<sup>82</sup> For 19th March meeting summary: <http://stopeugenics.org/files/2008/04/dh-letter-to-anna-middleton-response-to-meeting.pdf>. The Department of Health agreed to remove the express reference to deafness in the Explanatory Notes (p.100 para 66).

<sup>83</sup> House of Commons Science and Technology (n 43).

<sup>84</sup> Report on the Consultation on the review of the Human Fertilisation & Embryology Act 1990 (March 2006) <http://www.number10.gov.uk/Page16604>.

<sup>85</sup> *Ibid.*

<sup>86</sup> “Welfare of the child” and “best interests” are deemed to be analogous to physical harm.

<sup>87</sup> Wilkinson (n 10) p 20.

<sup>88</sup> Scott (n 23) p 317.



### 6.1 Substantial harm?

Harm can be conceptualised in both person-affecting and non person-affecting forms. These are considered in turn.

#### Person-affecting conceptions of harm

On applying the “substantial harm” requirement, the core issue is whether selecting for disability harms the disabled child. This would represent the strongest ground for a s.14(4)-type prohibition on selecting for disability.

The concept of harm is typically used in a comparative, person-affecting sense; harm attaches to a particular person, rather than to an intangible group of interests. Thus, it would be sensible to say that a child is harmed where its parents deliberately disabled it after birth. This is because that child would have not been disabled, but for the parents actions. Similarly, a child could claim to have been harmed where the parents interfered with the unaffected embryo, from which the child was created, to cause the disability.<sup>89</sup> In both examples, the child could point to a “non identity-affecting”<sup>90</sup> event that caused the disability and thus harmed her. That is, but for the parents’ actions, that child would not have been disabled.

In contrast, where a couple select for disability, the disabled child cannot point to a “non identity-affecting” event that caused her disability. These situations are not analogous with the previous examples because the couple have simply selected a disabled embryo that will become their disabled child. The couple could, of course, have selected a different, unaffected embryo but this “identity-affecting choice”<sup>91</sup> would have resulted in a different child; but for the couples’ actions that disabled child would not have been born. This is the inescapable conclusion of Derek Parfit’s famous “non-identity problem”;<sup>92</sup> a “systematic difficulty”<sup>93</sup> that shows choices made in the present can change the identity of future persons.

On this basis, the Government’s claim that selecting for disability harms the child, thus justifying s.14(4), relies on the premise that non-existence is preferable to life with a disability; non-existence being that child’s only alternative to life with a disability. This claim appears indefensible, and somewhat offensive, when made in relation to the vast majority of disabilities, such as achondroplasia and cystic fibrosis, which are, in the author’s opinion, compatible with lives worth living.<sup>94</sup>

The child would only have a valid claim to having been harmed where the disability is so severe that the child’s life is “not worth living.”<sup>95</sup> This would occur where the parents selected for Tay Sachs or Lesch-Nyhan syndrome, conditions that are, in the author’s opinion, so severe as to be not compatible with a life worth living.

To conclude, selecting for a disability that is compatible with a life worth living is a

<sup>89</sup> On the reasonable presumption that this alteration doesn’t affect his future identity.

<sup>90</sup> Wilkinson (n 10) p 3.

<sup>91</sup> *Ibid.*

<sup>92</sup> Parfit (n 27).

<sup>93</sup> Brock (n 27) p 269.

<sup>94</sup> Cf David Benatar *Better Never to Have Been: The Harm of Coming into Existence* (Oxford, 2006): Argues that the asymmetry between pleasure and pain leads all children to be harmed by being brought into the world.

<sup>95</sup> Allen Buchanan *et al From Chance to Choice* (Cambridge, 2000) p 224.



“morally neutral”<sup>96</sup> decision on person-affecting grounds, and certainly not one which causes “substantial harm” to the child. Thus, it cannot sensibly be said that selecting for disability is opposed to the “welfare” or “best interests” of the child, as the Government has suggested in relation to s.14(4). The non-identity problem also prevents the HFEA’s claim that selecting for disability breaches the s.13(5) welfare of the child requirement in the person affecting sense, except where the disability is not compatible with a life worth living<sup>97</sup> (see Part 3.2, *supra*).

### Non person-affecting conceptions of harm

The Government’s claim that selecting for disability is detrimental to the welfare of the child might be resurrected if harm, or welfare, is formulated in a non-person affecting sense. Derek Parfit’s Same Number Quality Claim (Q) states that: “If in either of two possible outcomes the same number of people would live, it would be worse if those who lived are worse off, or have a lower quality of life than those who would have lived.”<sup>98</sup> Q is a non person-affecting claim, which does not focus on the interest of a particular individual, but rather on the interests of groups of potential future people. Q intuitively appears correct; people should normally choose the option that leads to the higher quality of life.

In the selecting for disability scenario, there is a choice between a disabled and an unaffected embryo. On applying Q, the parents should choose the unaffected embryo, in the absence of competing reasons, as this embryo can reasonably be expected to have the best quality of life. The fact that the disabled embryo would still create a disabled child with a life worth living is irrelevant.

To explore this claim, it is useful to consider Julian Savulescu’s theory of Procreative Beneficence, which has reformulated Q in the bioethics context.<sup>99</sup> A recent formulation of this theory states:

If couples (or single reproducers) have decided to have a child, and selection is possible, then they have a significant moral reason to select the child, of the possible children they could have, whose life can be expected, in light of the relevant available information, to go best or at least not worse than any of the others.<sup>100</sup>

It is important to note that Procreative Beneficence is not absolute; “[T]he principle states, not what people invariably must do, but what they have a significant moral reason to do.”<sup>101</sup> This is because Procreative Beneficence is simply “a normative force” pushing the couple (or single reproducer) in the direction of selecting the best child, which is subject to competing normative reasons including the “welfare of the parents, of existing children, and of others, possible harm to others, and other moral constraints.”<sup>102</sup> Nevertheless, in the absence of such competing reasons, the couple ought, all things considered, select the most advantaged child,<sup>103</sup> which would be the unaffected embryo, because “it is not morally

<sup>96</sup> Gavaghan (n 15) p 85.

<sup>97</sup> Indeed, if selecting for disability did harm the child, s.14(4) wouldn’t be required, as it would already be prevented by s.13(5).

<sup>98</sup> Parfit (n 27) p 360.

<sup>99</sup> Savulescu (n 5); Julian Savulescu ‘Procreative Beneficence: Why We Should Select the Best Child?’ Bioethics Volume 15, Number 5/6 2001.

<sup>100</sup> Savulescu (n 5) p 274.

<sup>101</sup> Savulescu (n 5) p 278.

<sup>102</sup> *Ibid.*

<sup>103</sup> *Ibid.*

permissible to choose less than the best.”<sup>104</sup>

On this approach, selecting for disability is objectionable because “a worse-off future person is created when a better-off future person could have been created instead.”<sup>105</sup> Procreative beneficence *prima facie* appears to provide a sensible basis upon which a couple can be said to act contrary to morality when they select a disabled embryo.

The outstanding issue is whether the scope of Procreative Beneficence can be extended, beyond the moral sphere thus far envisaged, to show that selecting for disability causes the type of “substantial harm” required to justify Section 14(4). Such an extension appears to raise several problems.<sup>106</sup>

The first is that non-person affecting formulations of harm, by their very nature, rely on the “shared interests of groups of potential future persons”;<sup>107</sup> in the Condition X Scenario, this class would be the four embryos that could be implanted. The problem is that only the embryo selected for implantation will develop any interests at all; “to speak of the four as having a shared pool of interests flies in the face of this reality.”<sup>108</sup> Even if this problem is overcome, harm is only caused in an intangible sense, whereas a restriction on Procreative Liberty requires “substantial harm to the tangible interests of others.”<sup>109</sup>

Secondly, non-person affecting principles, such as Procreative Beneficence, only apply where the choices lead to the same number of people. In the achondroplastic scenario, however, it was seen that a parent might decide not to have a child if prohibited from selecting an achondroplastic child. This would be a “different number” situation.<sup>110</sup> Thus, the option of choosing the disabled child would not be morally wrong on non person-affecting grounds as the alternative choice would create no replacement; “If the numbers change or the burdens of substitutions are excessive, then the basis for arguing that there is a moral obligation to substitute vanishes... [as] one cannot be sure that one has maximized overall utility, which is the obligation at issue.”<sup>111</sup>

Thirdly, Savulescu accepts that Procreative Beneficence is not absolute, but rather can be outweighed by other normative forces. Thus, Procreative Beneficence would not justify a total prohibition on selecting for disability of the type imposed by s.14(4).<sup>112</sup>

The final, and perhaps most serious, problem is that Procreative Beneficence appears to prove too much. Procreative Beneficence would not be limited to restricting selecting for disability, but would further justify placing a legal duty on all parents to create the “best children.” As Wilkinson notes, this would “entail accepting that the case for prohibiting selecting disability is (*ceteris paribus*) equally a case for compulsory enhancement selection.”<sup>113</sup>

To conclude, Procreative Beneficence, and other non person-affecting principles might succeed in showing that selecting for disability is, in some instances, morally wrong rather

<sup>104</sup> *Ibid.*

<sup>105</sup> Wilkinson (n 10) p 13.

<sup>106</sup> For a strong critique of Procreative Beneficence: Robert Sparrow, “Procreative Beneficence, Obligations and Eugenics”, *Genomics, Society and Policy* 2007, Vol.3, No.3, p 43-59.

<sup>107</sup> Gavaghan (n 15) p 87.

<sup>108</sup> *Id.* p 88.

<sup>109</sup> Robertson (n 8) p 41.

<sup>110</sup> Parfit (n 27) p 356.

<sup>111</sup> John A Robertson ‘Procreative Liberty and Harm to Offspring in Assisted Reproduction’ *American Journal of Law & Medicine* (30) p 17; Parfit (n 27) p 370-371.

<sup>112</sup> Savulescu (n 5) p 278.

<sup>113</sup> Wilkinson (n 10) p 13. Suggests Prioritarianism to avoid this problem p 15-20.

than simply “morally neutral.”<sup>114</sup> These theories do not, however, provide a basis for a s.14(4)-type total prohibition on selecting for disability. Such a conclusion is affirmed by Savulescu, who recognises that Procreative Beneficence “is compatible, at the legal level, with enjoyment of a right to autonomy, including the right to make procreative choices which foreseeably and avoidably result in less than the best child.”<sup>115</sup>

## 6.2 Resources?

The second question is whether selecting for disability substantially harms the resources of others. The Government did not directly refer to resource allocation in its limited justification for s.14(4); an unsurprising omission given that it is a “highly sensitive” issue.<sup>116</sup> In this context, “resources” refers to the revenue that the Government generates from taxpayers that is used to fund the NHS and the social security system. It is submitted that such resources can sensibly be referred to as a “tangible interest,” and thus be a type of “substantial harm.”

Procreative Liberty would not require the Government to provide resources to allow a couple to select for disability. The couple would need to pay for the PGD themselves. This is because Procreative Liberty is a “negative right” (see 4.1, *supra*). Thus, the Government’s resources would not be implicated in the initial funding of such treatment, but it is necessary to consider how Governmental resources might be implicated following the birth of the disabled child.

Resources will clearly not be implicated where the parents are committed to raising the child and have the resources to ensure that it receives the appropriate level of care. In this situation, there is no basis on which to claim that selecting for disability has caused “substantial harm.” This view is perhaps shared by John A Robertson: “as long as parents who use these techniques are committed to rearing their child, they should be considered to be exercising or engaging in legitimate procreative activity.”<sup>117</sup>

If the parents lacked these traits, however, the Government would be morally obliged to provide assistance. In this situation, the State’s resources are clearly implicated. It might therefore be argued that allowing parents who do not have the resources or capabilities to rear the disabled child to select for disability causes “substantial harm” to the Government’s resources.<sup>118</sup>

On the other hand, the number of people likely to select for disability is so small that any impact on resources would be minimal. Nevertheless, the fact that resources are limited, and subject to numerous other competing demands, means that it is reasonable to describe such conduct as causing “substantial harm.”

Thus, it seems fair to require parents who want to select for disability to prove that they can raise the disabled child without any more assistance than the State would typically provide parents who had not selected for disability. This would likely represent a more significant burden in relation to serious disabilities such as cystic fibrosis, and a lesser

<sup>114</sup> For further moral concerns, the Child’s right to an Open Future: Davis (n 8) p 22-34, 63-67; Joel Feinberg ‘The Child’s Right to an Open Future’ in Will Aiken and Hugh La Fallette *Whose Child? Children’s Rights, Parental Authority, and State Power* (Littlefields, 1980) p 124-153.

<sup>115</sup> Savulescu (n 5) p 278.

<sup>116</sup> Scott (n 23) p 317.

<sup>117</sup> John A. Robertson. “Procreative Liberty in the Era of Genomics.” *American Journal of Law & Medicine*, 29 (2003) p 487.

<sup>118</sup> Alternatively, it could be argued that adequate resources are part of the internal constraints on Procreative Liberty: Robertson (n 111) 2004 p 21.

burden for disabilities such as achondroplasia and deafness.

## 7. Regulating Selecting for Disability: A New Approach?

### 7.1 Section 14(4): A critique

The HFE Act 2008 was introduced with the admirable intention of clarifying the law.<sup>119</sup> It is submitted, however, that legal clarity has been gained at the expense of legitimacy in the case of s.14(4).

The Government's failure to give effect to the force of Procreative Liberty is disappointing; particularly in light of responses received during the HFE Act 2008 review process. The House of Commons Science and Technology Committee alluded to Procreative Liberty when concluding that "the creation of a child with reduced life opportunities is not sufficient grounds for regulatory intervention."<sup>120</sup> Similarly, several respondents to the Department of Health's Public Consultation<sup>121</sup> stated that Procreative Liberty should protect the right to select for disability.<sup>122</sup> The Government's failure to expressly recognise, or respond to these views, negates the purpose of having a review process.<sup>123</sup> The HFEA can be similarly criticised for failing to implement Procreative Liberty, despite having expressly recognised its force.<sup>124</sup>

The opaqueness of the Government's decision-making process can also be criticised for failing to provide defensible arguments in support of s.14(4). The limited arguments that were raised, in relation to the welfare of the child, fail to justify a total prohibition on selecting for disability (Parts 6.1-6.2). In the absence of a transparent and defensible decision-making process, the Government leaves itself open to the charge of having "formalised the tyranny of the majority."<sup>125</sup>

Fundamentally, "legal and policy authorities, including licensing authorities such as the HFEA, need a stronger case than has yet been articulated for denying willing physicians and couple the freedom to use PGD" to select for disability in situations that met the requirements of Procreative Liberty.<sup>126</sup> The next section suggests that the HFEA and Government should adopt a more flexible approach to regulation in this area.

### 7.2 An alternative to s.14(4)

This article has argued that Procreative Liberty protects a citizen's right to select for disability where: (1) the internal constraint is met (see Part 4.2, *supra*); (2) the disability is compatible with a life worth living (Part 4.1); and (3) the parents have the resources, and inclination, to care for the disabled child (Part 6.3) (the three licensing conditions). It has been submitted that the three licensing conditions can be met in some situations,

<sup>119</sup> HFEA Listening (n 29) p 5.

<sup>120</sup> House of Commons Science and Technology (n 43) [145]. Recognised that "further debate" would be required in the "more challenging but unlikely scenario" that the parent sought "to select a child who would suffer obvious discomfort (rather than disadvantage), or worse."

<sup>121</sup> Review (n 16) [5.13] [5.20].

<sup>122</sup> British Fertility Society Response in *Report on the Consultation on the Review of the Human Fertilisation & Embryology Act 1990*. Prepared for the Department of Health (March 2006); p 36; Cf "Anyone deliberately inflicting an impairment on a foetus brought to term should be charged with 'grievous bodily harm'." Cambridge Genetics Knowledge Park Response ([http://www.dh.gov.uk/en/Consultations/Responsestoconsultations/DH\\_4132358](http://www.dh.gov.uk/en/Consultations/Responsestoconsultations/DH_4132358)).

<sup>123</sup> For example, Government Response to the Report from the House of Common Science and Technology Committee (*August 2005*): Did not respond to the Committee's commitments on selecting for disability.

<sup>124</sup> Harris (n 7).

<sup>125</sup> Harris (n 7) p 294.

<sup>126</sup> Robertson (n 9) p 101.

suggesting that a s.14(4)-type total prohibition on selecting for disability is unjustified.

Fundamentally, strict statutory provisions such as s.14(4) are not flexible enough to recognise the complexities in this area. The HFEA should be reinstated with the discretionary power that it used to enjoy in relation to licensing selecting for disability under the unamended HFE Act 1990. It is submitted that the HFEA Licensing Committee (HFEALC) should handle these applications, rather than delegating the discretion to the clinics. This approach is already utilised for other ‘controversial’ uses of PGD, such as HLA tissue typing and late onset conditions, and recognises the importance of ensuring selecting for disability is subject to strict, central controls. The HFEALC would be obliged to grant licenses where applications met the three licensing conditions.

If the HFEALC granted a licence, the applicants would have the right to select for disability. On a practical level, it is submitted that licenses are most likely to be granted in relation to achondroplasia and deafness; applications for serious conditions such as cystic fibrosis might struggle to satisfy the resource condition.

It is important to note that the HFEALC and clinics would not be able to refuse an application on the grounds that it harmed the child under s.13(5), unless the disability was not compatible with a life worth living. Thus, the HFEA would need to reinterpret s.13(5) in the Code of Practice to recognise that the welfare of the child is only diminished in these limited circumstances. Clinics would nevertheless be entitled to refuse to carry out treatment under the s.38(1) HFE Act 1990 conscientious objection clause; this would respect their right to autonomy.

Finally, the HFEALC should be required to closely monitor the impact of selecting for disability. This article has argued that “substantial harm” is not inevitably caused by selecting for disability in all cases. Nevertheless, it is possible that such selection could have an unacceptable affect on third parties, or create unforeseen societal problems, that might justify prohibition. There is no current empirical evidence, however, to suggest that such problems would arise.

To conclude, this approach to the regulation of selecting for disability would offer a flexible and sensible approach that recognizes the importance of Procreative Liberty.

## 8. Conclusion

Procreative Liberty does not require the Government to allow citizens free reign in the “genetic supermarket”<sup>127</sup> where reproductive technologies could be used for any purpose.<sup>128</sup> It does, however, entail a more flexible approach to regulation than the s.14(4) total prohibition on selecting for disability. This article has based criticisms of s.14(4) on the primacy of Procreative Liberty,<sup>129</sup> but similar conclusions would likely be reached under other regulatory models, such as Colin Farrelly’s “Reasonable Genetic Intervention Model.”<sup>130</sup> The Government’s reasoning for s.14(4) appears, in reality, to rest on “symbolic and dignitarian harms that cannot be directly tied to harm to offspring,”<sup>131</sup> rather than empirical evidence. A more flexible approach, such as that outlined in Part 7.2, is required.

<sup>127</sup> Robert Nozick, *Anarchy, State and Utopia* (1974).

<sup>128</sup> For a strong argument in favour of such an approach, see: Gavaghan (n 15).

<sup>129</sup> For general discussion of interaction between law and morality: Beauchamp and Childress, *Principles of Biomedical Ethics* (2009) p 9-10.

<sup>130</sup> Colin Farrelly, “Preimplantation Genetic Diagnosis, Reproductive Freedom and Deliberative Democracy,” *Journal of Medicine and Philosophy* 34 (2009) p 142.

<sup>131</sup> Robertson (n 111) p 18.



# IT'S JUST NOT CRICKET: EU COMPETITION LAW AND ITS APPLICATION TO SPORTING RULES

EOIN KEALY, BBLS  
KING'S COLLEGE LONDON  
EOINKEALY@GMAIL.COM

**ABSTRACT** This article explores the application of EU competition law to the rules and regulations of sporting associations. The article shows that, in the past, Articles 101 and 102 of the Treaty on the Functioning of the European Union were generally deemed inapplicable to such rules due to the specific nature of sport, which was considered to be distinct from other realms of economic activity to which the competition rules of the Treaty applied. This approach of the EU institutions has altered in recent years, and the judgment of the European Court of Justice (ECJ) in *Meca-Medina* confirmed that rules governing sporting activity are subject to the scrutiny of EU competition law. This article sketches the journey of the EU institutions from granting sporting federations autonomy in their creation of rules governing sport, to the recent, and long overdue, recognition of the Commission and the ECJ that sporting rules associated with an economic activity cannot be deemed immune to the application of competition law. The article outlines a number of current sporting practices and regulations which may be suitable for challenge under EU competition rules. However, the article proposes that sporting rules may still be treated in a less rigorous manner under competition law in the future, although more challenges to such rules under EU competition law is inevitable.

## Introduction<sup>1</sup>

Competition lawyers and stakeholders in the sporting sphere are lamenting the recent removal of the *Oulmers* case<sup>2</sup> from the pending case register of the European Court of

<sup>1</sup> Note that the Treaty on the Functioning of the European Union ('TFEU') entered into force on 1 December 2009, changing the numbering of many of the provisions of the Treaty Establishing the European Community ('EC Treaty' or 'EC'). Where possible, I will refer to the new numbering in the TFEU. However, much of the authority cited and quoted in this article will refer to the numbering under the EC Treaty. Article 1 TFEU has also replaced 'The European Community' with 'The European Union', therefore I will refer to 'EU law' rather than 'Community law' whenever possible. The Court of First Instance has been renamed the 'General Court' by the TFEU. Again, I will refer to the new title where possible.

<sup>2</sup> Case C-243/06 *SA Sporting du Pays de Charleroi, G-14 Groupement des clubs de football européens v Fédération internationale de football association (FIFA)* [2006] OJ C212/18; Notice of the removal of the case from the register is available at [2009] OJ C69/56.



Justice (ECJ). The judgment of the ECJ was eagerly anticipated in the expectation that it would provide much desired clarity in the treatment of sporting rules by the competition law provisions of the Treaty on the Functioning of the European Union (TFEU).<sup>3</sup> The issue in *Oulmers* on which the ECJ was requested to give a preliminary ruling pursuant to what is now Article 267 TFEU<sup>4</sup> was relating to the validity of FIFA rules obliging football clubs to release their players for international matches, while receiving no payment or compensation for their absence from their place of employment. The rules also provided that it was the obligation of the club to insure the player against the risk that the player would be injured while representing his country.

In this case, the football player, Oulmers, was injured while on international duty with Morocco. As per the FIFA rules, Charleroi did not receive any compensation for his long-term absence while continuing to pay his wages while he was injured. Charleroi brought a case to the Belgian courts seeking damages from FIFA for a breach of Article 82 EC. The outcome of this case was clearly of great import for football clubs, especially the most wealthy and successful clubs in Europe, who naturally employ the greatest proportion of internationally representative footballers. Consequently, the G-14 grouping of major European clubs intervened in the case on behalf of Charleroi. National associations' interests lay in preserving the status quo; over fifty continental and national associations intervened on behalf of FIFA.

The Tribunal de Commerce in Charleroi referred the following question to the ECJ:

Do the obligations on clubs and football players having employment contracts with those clubs imposed by the provisions of FIFA's statutes and regulations providing for the obligatory release of players to national federations without compensation and the unilateral and binding determination of the coordinated international match calendar constitute unlawful restrictions of competition or abuses of a dominant position or obstacles to the exercise of the fundamental freedoms conferred by the EC Treaty and are they therefore contrary to Articles 81 and 82 of the Treaty or to any other provision of Community law, particularly Articles 39 and 49 of the Treaty?<sup>5</sup>

The judgment of the ECJ was anticipated as an opportunity to clarify the scope of the application of EU competition law to the rules of sporting bodies following the landmark ECJ judgment in *Meca-Medina*.<sup>6</sup> In this article I propose to outline the historical context of the relationship between EU competition law and sporting rules, how the dynamic has altered as the nature of sporting activity has changed, and where the future lies in terms of sporting rules or practices which may be vulnerable to the scrutiny of competition law. I will also examine whether the current manner in which competition law and sporting regulations interact is satisfactory for the purposes of legal certainty.

## The Treaties and Sport

The first point to note is that sporting activity is not in itself subject to the competition rules of the TFEU.<sup>7</sup> It is settled law that sport is subject to EU law only in so far as it

<sup>3</sup> Articles 101 and 102 TFEU, formerly Articles 81 and 82 EC.

<sup>4</sup> Formerly Article 234 EC.

<sup>5</sup> *Oulmers* (n 2).

<sup>6</sup> Case C-519/04P *Meca-Medina and Majcen v Commission* [2006] ECR I-6991, [2006] 5 CMLR 18.

<sup>7</sup> Bellamy & Child, *European Community Law of Competition* (6th ed. Oxford University Press, Oxford, 2009) [5.156].

constitutes an 'economic activity' within the meaning of Article 2 EC.<sup>8</sup> Sporting federations are therefore able to avoid the application of EU law to its rules and regulations where it can be shown that the sport to which they apply does not constitute 'economic activity'. This gives the federations a certain degree of autonomy in their rulemaking, which has been recognised and reinforced at an institutional level on a number of occasions. In the Declaration on Sport (Declaration 29) of the Treaty of Amsterdam,<sup>9</sup> the Intergovernmental Conference on the Treaty of Amsterdam called on the bodies of the European Union to listen to sports associations when important questions affecting sport are at issue and to give special consideration to the particular characteristics of amateur sport. A similar set of ideals are set out in a Declaration attached to the Presidency Conclusions of the Nice European Council.<sup>10</sup> The 'soft law' form of the Declarations betrays a fundamental uncertainty; Declarations are non-binding instruments. In the absence of a Treaty provision on sport, any action by the EU was necessarily limited to instruments such as Declarations, as no provision for action in the Treaties means that there was no legal basis for EU action in the sporting arena. However, the TFEU entered into force on 1 December 2009, and Article 165 TFEU contains express competences for the EU to contribute to the promotion of European sporting issues by adopting 'incentive measures'<sup>11</sup> and recommendations, but expressly excludes harmonisation.

While a specific basis for EU action in sport is welcome, the competence of the Union to act remains limited. Van den Bogaert and Vermeersch<sup>12</sup> opine that "[t]he total exclusion of harmonising measures is regrettable" and that the new article is "...of a mere symbolic nature."

The absence of a mention of sport in the EC Treaty does not mean that the provisions of Community law did not, and will not in the future, apply to sporting rules. As mentioned above, sport comes within the scope of EU law in so far as it constitutes an economic activity. Articles 101 and 102 TFEU, as well as the free movement provisions of the TFEU, will apply when a sport constitutes economic activity. Clearly, sport in Europe is driven, to a large degree, by commerce. Sport in the EU generated value added of approximately €407 billion in 2004, amounting to 3.7% of EU GDP and employment for 15 million people or 5.4% of the labour force.<sup>13</sup> While sport may aspire to social and cultural ideals, it is a major industry and thus must take heed of EU competition law.

### 'Purely Sporting' Rules

The EU institutions have recognised the "specificity of sport" in that its rules, activities and structures are specific to the sporting domain and are necessary for its proper

<sup>8</sup> Case 36/74 *Walrave and Koch v Union Cycliste Internationale* [1974] ECR 1405; Case 13/76 *Donà v Mantero* [1976] ECR 1333; Joined Cases C-51/96 & C-191/97 *Christelle Delège v Ligue Francophone de Judo et Disciplines Associées ASBL, Ligue Belge de Judo ASBL, Union Européenne de Judo; Christelle Delège v Ligue Francophone de Judo et Disciplines Associées ASBL, Ligue Belge de Judo ASBL, François Pacquée* [2000] ECR I-2549; Case C-176/96 *Jyri Lehtonen and Castors Canada Dry Namur-Braine ASBL v Fédération Royale Belge des Sociétés de Basket-ball ASBL (FRBSB)* [2000] ECR I-2681; *Meca-Medina* (n 6).

<sup>9</sup> Declaration No 29 in the Final Act of the Intergovernmental Conference on the treaty of Amsterdam signed on 2 October 1997 [1997] OJ C340/136.

<sup>10</sup> "Declaration on the specific characteristics of sport and its social function in Europe, of which account should be taken in implementing common policies," Nice European Council (7, 8 and 9 December 2000), Presidency Conclusions (paragraph 54 and Annex IV).

<sup>11</sup> Article 165(4) TFEU.

<sup>12</sup> Stefaan Van den Bogaert and An Vermeersch, "Sport and the EC Treaty: a tale of uneasy bedfellows?" (2006) 31(6) *EL Rev* 821, 838.

<sup>13</sup> Commission (EC), "White Paper on Sport" (White Paper) COM (2007) 391 final, 11 July 2007, p 11.

functioning.<sup>14</sup> For example, it is in this context that the ECJ ruled in *Walrave*<sup>15</sup> that when a sporting rule is of ‘purely sporting interest’ it will not be subject to Community law. In this case, a rule restricting the nationality of an athlete on a national cycling team was deemed to be ‘purely sporting’ and therefore non-economic and outside the scope of Community law. The ECJ again followed this approach in *Donà*,<sup>16</sup> where again a rule confining a place in a national team to nationals was upheld as being purely sporting. Cases involving rules applying to national teams can be contrasted with the *Bosman* case.<sup>17</sup> In *Bosman*, UEFA rules permitted each national football association to limit to three the number of foreign players a club may play in a match in their national championship, plus two players who had played in the country of the national association for an uninterrupted period of five years (the ‘3+2 rule’). A separate FIFA transfer rule requiring payment of international end-of-contract transfer fees within the EU in respect of players who are nationals of an EU Member State was also scrutinised by the Court. The ECJ held that the rules breached the free movement of workers provision in Article 39 EC (now Article 45 TFEU). Article 39 precluded restrictions by sporting associations on the number of nationals from EU Member States taking part in international or national club competitions.<sup>18</sup> The ECJ decided the case on Article 39, but both the Commission and Advocate General (AG) Lenz were of the opinion that rules restricting the employment of foreign players breached Article 81(1) EC as they limited the possibilities for individual clubs to compete with each other by employing players. The *Bosman* judgment did not affect the legitimacy of national association rules permitting only nationals of the Member State concerned to represent the national team. The nationality rule in *Bosman* was not of ‘purely sporting’ interest so as to escape the scope of Community law. With regard to the FIFA transfer rule which prevented an out-of-contract player from being employed by another football club unless an agreed transfer fee was paid, the ECJ again ruled that the regulation was contrary to Article 39 EC, without examining the rule under Article 81. However, AG Lenz opined that the rule did contravene Article 81. According to the Advocate General the transfer rules:

... replace the normal system of supply and demand by a uniform machinery which leads to the existing competition situation being preserved and the clubs being deprived of the possibility of making use of the chances, with respect to the engagement of players, which would be available to them under normal competitive conditions.<sup>19</sup>

AG Lenz did consider that a transfer fee in such cases may be lawful if, firstly, the fee was limited to the amount spent by the previous club or clubs on the players training, and, secondly, a transfer fee would only be permitted in the case of a first change of clubs where the previous club had trained the player.<sup>20</sup> This would be in order so that the training club would benefit from its investment in the player.

*Bosman* represented a watershed in how the EU viewed its relationship with sporting regulations. While the institutions had previously adopted a ‘hands-off’ approach to the

<sup>14</sup> Commission (EC), "Report from the European Commission to the European Council with a view to safeguarding current sports structures and maintaining the social function of sport within the Community framework ("Helsinki Report on Sport") COM (1999) 644, 10 December 1999; "White Paper on Sport," *ibid* p 13.

<sup>15</sup> *Walrave and Koch* (n 8).

<sup>16</sup> *Donà v Mantero* (n 8).

<sup>17</sup> Case C-415/93 *URBSFA v Bosman* [1995] ECR I-4921.

<sup>18</sup> *ibid* at [137].

<sup>19</sup> *ibid* AG Lenz at [262].

<sup>20</sup> *ibid* AG Lenz at [239].

regulations of sports associations, recognising the specific characteristics of sport and the autonomy of sporting regulatory bodies, the ECJ in *Bosman* demonstrated that sporting rules made by autonomous associations will not, in all circumstances, escape the application of EU law. As Szyszczak points out, the *Bosman* judgment contained inherent weaknesses:

... it emphasised the incompatibility of the FIFA football transfer regime with the fundamental economic provisions of the EC Treaty but provided no guidance on how to resolve the issues, particularly when competition law provisions were applied to sport-related activities. The Court acknowledged, but refused to provide a clear framework on how to address “(...) the difficulty of severing economic aspects from the sporting aspects of football.”<sup>21</sup>

*Bosman* highlighted the dichotomy between previous cases like *Walrave*, where rules were seen as being of ‘purely sporting interest,’ and the commercial reality of modern sport. Sport is big business, and, as a result, the laws regulating it have major economic effects on all of the various investors, media undertakings, supporters and participants. The increased commercialisation of sport has made it less plausible to argue that a sporting rule alleged to breach EU competition law, or other Treaty provisions, is purely sporting and thus evades the scope of EU law, when it has clear economic effects on sporting markets.

### From *Bosman* to *Meca-Medina*

Post-*Bosman* and before the important *Meca-Medina*<sup>22</sup> judgment of the ECJ, a number of cases involving sporting rules allegedly breaching Articles 81 and/or 82 EC came before the Commission and the Community courts. These cases were not solely based on nationality requirements regarding selection for teams, but represented a broad spectrum of sporting regulations, encompassing rules regarding football players’ agents to transfer rules, club ownership regulations to transfer windows.

As a result of the *Bosman* judgment, FIFA amended its transfer rules, but following a number of complaints, the Commission issued a statement of objections to FIFA regarding remaining situations where transfer fees still had to be paid for players whose contracts had expired.<sup>23</sup> The Commission also objected to transfer rules regarding players who were still under contract. It claimed a breach of Article 81 EC and free movement principles in that there was an effective ban on the unilateral breach of contract by a player. Under the rules, if a player breached his contract with his employing club and paid the requisite amount of damages for the breach, he still could not be registered to play with a new club in a different country. The Commission also claimed that the FIFA rules that provided for an ‘agreed’ transfer whereby the buying club, selling club and the player all agreed on a fee to be paid, but the fee was not required to bear any relation to the investment the selling club had made in the players training and development were a breach of Article 81 EC and the free movement provisions. Negotiations between the Commission and FIFA proved successful, and the Commission declared itself satisfied with a new set of FIFA rules on the international transfer of players.<sup>24</sup> The agreed principles set out measures establishing

<sup>21</sup> Erika Szyszczak, ‘Competition and sport’ (2007) 32(1) EL Rev 95, 97, citing *Bosman* (n 17, at [76]).

<sup>22</sup> *Meca-Medina* (n 6).[23] For a list of these situations see Adam Lewis and Brian Kennelly, Adam Lewis and Jonathan Taylor (eds), *Sport: Law and Practice* (2nd ed. Tottel Publishing, London 2008) Chapter B2, [2.164].

<sup>23</sup> For a list of these situations see Adam Lewis and Brian Kennelly, Adam Lewis and Jonathan Taylor (eds), *Sport: Law and Practice* (2nd ed. Tottel Publishing, London 2008) Chapter B2, [2.164].

<sup>24</sup> ‘Outcome of discussions between the Commission and FIFA/UEFA on FIFA Regulations on international football transfers’; Commission press release IP/01/314 of 6 March 2001.

defined transfer periods per season (transfer windows), measures providing for training compensation for young players under twenty three years of age and a solidarity mechanism providing for the redistribution of income from a transfer fee to clubs involved in training and developing a player, measures specifying duration of contracts and when breaches of contract are possible (and sanctions for a breach).<sup>25</sup> The Commission did not take a decision under Article 81(3) EC to exempt the new transfer rules. Instead, it discontinued its previous case against FIFA, which means that a challenge to the transfer rules may still be taken by an interested party.<sup>26</sup>

In *Mouscron*,<sup>27</sup> the *Communaute Urbaine de Lille* made a complaint to the Commission under Article 82 EC with regard to the “at home and away from home” rule which applies to UEFA competitions, whereby the home game of a tie over two legs is played at the home ground of the relevant club. The Commission deemed this to be a sporting rule that did not fall within the scope of Articles 81 and 82 EC. The rule was deemed indispensable for the organisation of national and international competitions in order to ensure equality of chances between clubs, and the rule did not go beyond what was necessary to achieve this objective. In *ENIC/UEFA*,<sup>28</sup> ENIC (a company which owned stakes in six professional football clubs in a number of Member States) complained to the Commission that a UEFA rule preventing more than one club which was directly or indirectly controlled by the same entity or managed by the same person from participating in a UEFA club competition breached Article 81(1) EC. The Commission concluded that the rule did not have the object of restricting competition, rather its object was to protect the integrity of UEFA tournaments by avoiding any perception that the outcome of a match was not uncertain, and the rule did not go beyond what was necessary to ensure the legitimate aim of protecting the uncertainty of the results of games in the interests of protecting credible competition.

*Mouscron* and *ENIC* are examples of the Commission recognising the ‘specificity of sport’,<sup>29</sup> in that there is a requirement that the result must be uncertain, and in order to achieve this objective there must be a degree of equality in competitions. This can be contrasted with other sectors where competition results in the elimination of inefficient firms from the market. Sport is unique in that it is not in the interests of competitors that weaker teams are eliminated. The Commission recognises that the unique nature of sport may affect the analysis of organisational sporting rules under EU law.

The Court of Justice further examined selection rules regarding national teams in *Deliège*,<sup>30</sup> where a judoka claimed that the system for selecting judokas for national tournaments breached her right to provide services and her professional freedom. European Judo Union rules stated that only national federations could enter athletes in certain tournaments, and only seven men and seven women could be entered for an event. *Deliège* was claiming, in essence, that by not being selected for the national team, her freedom to provide services was being restricted, and that the rules were anti-competitive.

<sup>25</sup> Lewis and Taylor (n 23) Chapter B2, [2.173]; Commission (EC), ‘The EU and Sport: Background and context. Accompanying document to the White Paper on Sport’ (Commission staff working document) COM (2007) 395, 11 July 2007, Annex 1, [2.2.2.4.2].

<sup>26</sup> Lewis and Taylor (n 23) Chapter B2, [2.175].

<sup>27</sup> *C.U. de Lille/ UEFA (Mouscron)* (Case 36851). Decision not published; see Commission press release IP/99/965 of 9 December 1999.

<sup>28</sup> *ENIC/UEFA* (Case COMP/37.806). Decision of 25 June 2002; see Commission press release IP/02/942 of 27 June 2002.

<sup>29</sup> See n 14, above.

<sup>30</sup> *Christelle Deliège v Ligue Francophone de Judo et Disciplines Associées ASBL* (n 8).



The ECJ found that, although the selection rules limited the number of judokas who could participate in a competition, the limitation was inherent in the conduct of an international sports event, which necessarily involved certain selection rules and limits on the number of participants. Such rules could not, in themselves, constitute an unlawful restriction on the freedom to provide services,<sup>31</sup> and was not discriminatory. The ECJ also acknowledged the 'pyramid' structure of the sport of judo, in that it was for sports federations to lay down appropriate rules and make selections accordingly, and that "...the delegation of such a task to the national federations, which normally have the necessary knowledge and experience, is the arrangement adopted in most sporting disciplines, which is based in principle on the existence of a federation in each country".<sup>32</sup>

In *Lehtonen*,<sup>33</sup> the ECJ was invited to determine the validity of transfer deadlines under competition rules and free movement of workers provisions. Transfer deadlines prevent clubs from registering (i.e. buying) players after a certain point in the season, so as to prevent teams from purchasing players at critical points in the season thereby distorting the integrity of the competition. This type of rule is again a function of the 'specificity of sport,' where a restriction on competition which would be inconceivable in another sector is fundamental to the nature of sporting competition. In this case, the court held that transfer deadlines were a de facto restriction on the free movement of workers, but could possibly be justified on non-economic, purely sporting grounds. Setting transfer deadlines "...may meet the objective of ensuring the regularity of sporting competitions"<sup>34</sup> in that late transfers could substantially alter the strength of a team in the tournament, thus calling into question the proper functioning of the competition. The court noted that such a restriction must be proportionate, i.e. the measure must not go beyond what is necessary to achieve the aim of ensuring the proper functioning of the tournament. In this case, the transfer deadline was deemed by the ECJ to be disproportionate in that it discriminated between players being transferred from a federation outside the European zone and players being transferred from a federation inside the European zone.

It should be noted that neither *Deliège* nor *Lehtonen* was decided under the competition rules of the Treaty, as the court in both cases deemed that it had not received sufficient information from the referring national courts on the factual and legal elements of the dispute, therefore it could not pronounce on the application of Articles 81 and 82 EC to the case. Turner-Kerr and Bell<sup>35</sup> declare this to be "regrettable", but deduce two principles from the judgments. Firstly, rules which are inherent in the conduct and/or organisation of sporting events do not, in themselves, infringe EU law, and, secondly, in relation to such rules, it is for the sports federation in question to decide what the appropriate measures are.<sup>36</sup>

The Court of First Instance (CFI) (now renamed the General Court by the TFEU) ruled on FIFA regulations concerning players' agents in *Piau*.<sup>37</sup> It was alleged that rules stating

<sup>31</sup> *ibid* at [64].

<sup>32</sup> *ibid* at [68]. For further discussion on the 'pyramid' structure as a characteristic of the 'European Model of Sport' see the Helsinki Report on Sport (n 14); Section 4, "The EU and Sport: Background and context. Accompanying document to the White Paper on Sport" (Commission staff working document) (n 25); and Stephen Weatherill, "Is the Pyramid Compatible with EC Law?" (2005) 3-4 ISLJ 3.

<sup>33</sup> *Lehtonen v FRBSB* (n 8).

<sup>34</sup> *ibid* at [53].

<sup>35</sup> Peter Turner-Kerr and Alasdair Bell, "The place of sport within the rules of Community law: clarification from the ECJ? The *Deliège* and *Lehtonen* cases" (2002) 23(5) ECLR 256.

<sup>36</sup> *ibid* p 260.

<sup>37</sup> Case T-193/02 *Laurent Piau v Commission* [2005] ECR II-209 (upheld by the ECJ in Case C-171/05P [2006] ECR I-37).



that a football player's contract was valid only if the agent involved in its conclusion had a licence issued by the national football association was a breach of Articles 81 and 82 EC. The rules stated that licensed agents must pass an interview, have an impeccable reputation and deposit a bank guarantee. FIFA amended the rules as a result of the Commission's investigation; therefore the complaint was rejected by the Commission. On appeal, the CFI stated that the activities of a football agent clearly do not pursue a purely sporting interest, and that any attempt by a private association like FIFA to restrict the economic pursuit of football agency must be questioned as to its legitimacy. Rules adopted by FIFA in this regard would be subject to EU law. The CFI did acknowledge that the profession needed to be supervised by some entity, which, in the absence of national laws or public bodies providing for supervision and the lack of internal self-regulation among agents, did not exist. The CFI upheld the Commission's conclusion that the rules did not produce anti-competitive effects under Article 81 EC, and even if such effects existed, they could be exempted under Article 81(3) EC.

*Piau* was a sign that the CFI was prepared to examine the legitimacy of a private body like FIFA exercising quasi-public discretion to limit fundamental freedoms guaranteed by the Treaty. Van den Bogaert and Vermeersch find fault in the CFI's reasoning, in that the finding that there were no national rules and no collective organisation of players' agents was inaccurate.<sup>38</sup> However, the ruling was a shot across the bows of FIFA and other practical monopoly sports federations that the EU institutions were willing to examine their regulations under competition law.

### Meca-Medina

It is fair to say that the ECJ's judgment in *Meca-Medina*<sup>39</sup> has fundamentally altered the manner in which sporting rules interact with Articles 101 and 102 TFEU. This was the first time that the ECJ had pronounced on the application of these Treaty provisions to an organisational sporting rule.<sup>40</sup> In previous cases where Articles 81 and/or 82 EC (now Articles 101 and 102 TFEU) were pleaded, the ECJ had only ruled on the basis of free movement provisions of the Treaty,<sup>41</sup> although various Advocate-Generals had included analysis of the competition law claims in their opinions.<sup>42</sup>

In *Meca-Medina*, two long distance swimmers representing Spain and Slovenia tested positive for nandrolone and were suspended for four years (reduced to two years on appeal) for breaching the anti-doping rules set by the International Olympic Committee (IOC) which are implemented in swimming by FINA, the International Swimming Federation. The two athletes alleged that the anti-doping regulations were a breach of Article 81 EC, that the establishment of a threshold amount of nandrolone permissible in an athlete's body was a concerted practice between FINA, the IOC and the laboratories carrying out the testing, and that the dispute settlement system established by the IOC was anti-competitive.

It is worth analysing the contrasting approaches taken by the Commission, the CFI and

<sup>38</sup> Van den Bogaert and Vermeersch (n 12), p 835.

<sup>39</sup> *Meca-Medina* (n 6).

<sup>40</sup> The Commission points out at footnote 99 in its staff working document accompanying the White Paper on Sport (see n 25) that the GC's judgment in *Piau* concerned a sporting rule adopted in relation to football agency (an activity ancillary to sport) and not relation to the sporting activity itself (football).

<sup>41</sup> *Bosman* (n 17); *Deliège* (n 8); *Lehtonen* (n 8).

<sup>42</sup> AG Lenz in *Bosman* (n 17); AG Cosmas in *Deliège* (n 8).

the ECJ to the case. The Commission relied on the ECJ's judgment in *Wouters*<sup>43</sup> in order to reject the complaint. In *Wouters*, the ECJ ruled that not every agreement, decision or concerted practice which restricts the freedom of action of undertakings falls within the scope of Article 81(1) EC. The ECJ stated that:

... account must... be taken of the overall context in which the decision... was taken or produces its effects. More particularly, account must be taken of its objectives... [i]t has then to be considered whether the consequential effects restrictive of competition are inherent in the pursuit of those objectives.<sup>44</sup>

The ECJ ruled that the regulation in question (a rule of the Netherlands Bar stating that lawyers could not enter into partnership with accountants) could "...reasonably be considered to be necessary in order to ensure the proper practice of the legal profession, as it is organised in the Member State concerned."<sup>45</sup>

*Wouters* was a controversial judgment, and attracted much comment,<sup>46</sup> in that it contrasted with previous case-law of the court which seemed to suggest that a restriction on competition could only escape the scope of Article 81(1) EC on the basis of an economic or commercial necessity.<sup>47</sup> In *Wouters*, the ECJ weighed non-competition objectives (in this case, the integrity of the system of justice in the public interest) against the restriction of competition (preventing lawyers from forming multi-disciplinary partnerships with accountants) and decided that the non-competition public interest objectives trumped the restriction on competition inherent in the regulation. This had the purpose of allowing the rule to escape Article 81(1) EC, and therefore there was no requirement for the authors of the rule to justify it under the stringent requirements of Article 81(3) EC.

The Commission explicitly relied on *Wouters* in rejecting the complaint against the anti-doping rules.<sup>48</sup> It decided that the anti-doping rules did not have the object of restricting competition. The rules may have had that *effect*, but such a restriction on the athletes' freedom to compete was not necessarily a breach of Article 81(1) EC as it may be inherent in the organisation and proper conduct of sporting competition. The Commission concluded that the anti-doping rules were intimately linked to the proper conduct of sporting competition, they were necessary to combat doping effectively and the limitation of the athletes' freedom did not go beyond what was necessary to attain that objective.

On appeal, the CFI also rejected the claim of the complainants, but using a different method.<sup>49</sup> It used the *Walrave* case to maintain that sport is subject to EU law only in so far as it constitutes an economic activity within the meaning of Article 2 EC,<sup>50</sup> and that the rules of competition will not apply to rules of purely sporting interest.<sup>51</sup> The court concluded that the anti-doping rules were of purely sporting interest, designed to preserve

<sup>43</sup> Case C-309/99 *Wouters v Algemene Raad van de Nederlandse Orde van Advocaten (Raad van de Balies van de Europese Gemeenschap intervening)* [2002] 4 CMLR 913.

<sup>44</sup> *ibid* at [97].

<sup>45</sup> *ibid* at [107].

<sup>46</sup> See Giorgio Monti, 'Article 81 EC and Public Policy' (2002) 39 CML Rev 1057; Rosemary O'Loughlin, 'EC Competition Rules and Free Movement Rules: An Examination of the Parallels and their Furtherance by the ECJ *Wouters* Decision' (2003) 24 ECLR 62.

<sup>47</sup> E.g. Case 56/65 *Société Technique Minière v Maschinenbau Ulm* [1966] ECR 235; Case 42/84 *Remia BV and Verenigde Bedrijven and Nutricia v Commission* [1985] ECR 2545.

<sup>48</sup> Case COMP/38.158. Text of Decision available in French at <http://ec.europa.eu/competition/antitrust/cases/decisions/38158/fr.pdf> accessed 23 December 2009.

<sup>49</sup> Case T-313/02 *David Meca-Medina and Igor Majcen v Commission* [2004] ECR II-3291.

<sup>50</sup> *ibid* at [37].

<sup>51</sup> *ibid* at [41].

“noble competition,”<sup>52</sup> and therefore escaped the scope of the Treaty rules on competition. The CFI’s judgment attracted much criticism for the manner in which it arrived at its conclusion.<sup>53</sup> The ruling that anti-doping rules, with clear economic effects for the athletes concerned (depriving them of professional fees and sponsorship) and for the sporting federations (drug free competitions are likely to be more financially lucrative in attracting sponsorship and television monies) are rules of purely sporting interest is an impractical and intellectually supine method of validating the regulations. In effect, it would mean that regardless of how unreasonable and draconian, or, indeed, unduly lenient sports federations’ anti-doping rules were, they would always escape the application of the Treaty. The CFI, when presented with the *Wouters* judgment as a way of permitting the rules, instead explored “intellectually murky alleyways”<sup>54</sup> so as not to interfere with the autonomy of the IOC and FINA in applying anti-doping rules.

In an important judgment on appeal, the ECJ rejected the CFI’s approach.<sup>55</sup> The court held that:

...the mere fact that a rule is purely sporting in nature does not have the effect of removing from the scope of the Treaty the person engaging in the activity governed by that rule or the body which has laid it down.<sup>56</sup>

This means that there is no sporting rule which can escape the Treaty rules per se, when the sport constitutes an economic activity within the meaning of Article 2 EC. The rules relating to such a sport need to satisfy the Treaty provisions on competition and free movement. The ECJ then applied *Wouters* to the anti-doping rules. It looked at the overall context in which the rules were adopted, and the general objective of the rules, which was to combat doping in order for competitive sport to be conducted fairly, and that it included the need to safeguard equal chances for athletes, athletes’ health, the integrity and objectivity of competitive sport and ethical values in sport.<sup>57</sup> The ECJ held that the penalties and their effect on athletes’ freedom of action were, in principle, inherent in the anti-doping rules and the legitimate objectives of those rules. The court did hold that the penalties imposed by the rules were capable of producing adverse effects on competition if they were unjustified and resulted in the unwarranted exclusion of an athlete from a competition. Therefore, the court held that, to escape the scope of Article 81(1) EC, the restrictions imposed by the anti-doping rules “...must be limited to what is necessary to ensure the proper conduct of competitive sport.”<sup>58</sup> In effect, the penalties must be proportionate. The ECJ, for obvious reasons, was reluctant to substitute its view on whether the level of nandrolone permissible in an athletes’ body was excessively low. The court is not best placed to make this type of scientific judgement.

### Implications of Meca-Medina

The judgment of the ECJ was consistent with the Commission’s analysis. Indeed, the

<sup>52</sup> *ibid* at [49], citing AG Cosmas’ Opinion in *Delège* at [50] and [74] (n 8).

<sup>53</sup> See Stephen Weatherill, “Anti-doping rules and EC law” (2005) 26(7) ECLR 416; Julian Gregory, “From Rio to Meca: another step on the winding road of competition law and sport” (2005) 64(1) CLJ 51.

<sup>54</sup> Stephen Weatherill, “Anti-doping revisited – the demise of the rule of ‘purely sporting interest’?” (2006) 27(12) ECLR 645, 648.

<sup>55</sup> *Meca-Medina* (n 6).

<sup>56</sup> *ibid* at [27].

<sup>57</sup> *ibid* at [43].

<sup>58</sup> *ibid* at [47].

Commission had previously applied the *Wouters* criteria in *Mouscron*<sup>59</sup> and *ENIC*.<sup>60</sup> The significance of *Meca-Medina* is that the *Wouters* criteria can now be applied to all sporting rules which have an economic effect. Given the commercial nature of many sports, it is difficult to think of rules made by sporting federations which do not have an economic impact. Even sporting rules made by federations governing amateur sports may have an economic impact, and come under the scrutiny of EU competition law. For example, the Gaelic Athletic Association (GAA) governs purely amateur Gaelic games on the island of Ireland, and in other areas such as London and New York, where there are significant Irish Diasporas. Elite Gaelic games are big business however, evidenced by the amounts paid by Irish television stations for the broadcasting rights for championship games.<sup>61</sup> Clearly, rules made by the GAA affecting these games (such as disciplinary regulations affecting the ability of marquee players to play in the championships) would have an economic effect on the value of these television rights, and on the value of games to the GAA itself in terms of gate receipts and partnership deals with sponsors. Indeed, the ECJ in *Deliège*<sup>62</sup> previously stated that the fact that a sports federation classifies its members as amateur athletes does not mean that the athletes do not carry on economic activities within the meaning of Article 2 EC. Another sporting rule which may attract the scrutiny of EU competition law in the future is FIFA's, or indeed other football federations', refusal to permit goal-line video technology in order to clarify contentious incidents in matches leading to goals, such as offside decisions, whether the ball crossed the line or foul play leading to goals. There are clear adverse economic impacts on a team which concedes a goal in this fashion, and the rules deny them an opportunity to challenge the goal. Controversial incidents such as Luis Garcia's winning goal for Liverpool against Chelsea in the UEFA Champions League semi-final in 2005 and Thierry Henry's handball leading to William Gallas' winner in the 2010 FIFA World Cup play-off game against Ireland led to significant loss of revenues for Chelsea in not qualifying for the final and, potentially, the Football Association of Ireland in not qualifying for the World Cup Finals. The practice clearly affects competition by distorting the results of games. FIFA could possibly claim that the restriction is inherent in the rules, and the objectives may be legitimate (e.g. ensuring the seamless running of the match or protecting the integrity of the referee), but can it plausibly be claimed that the restriction is proportionate to the objectives? In this author's view, an absolute ban on such video technology may be disproportionate when it is considered that technology is used successfully in elite rugby and tennis. A compromise solution involving limited use of video technology would stand a better chance of achieving proportionality under Articles 101 and/or 102 TFEU.

Weatherill<sup>63</sup> believes that the judgment eliminates the 'purely sporting' rule enunciated in *Walrave*, whereby sporting federations had autonomy to apply rules which would not fall to be analysed under the Treaty. Szyszczak believes that approaching each sporting rule on a case-by-case basis, as advocated by *Meca-Medina*, will likely lead to more litigation on the extent of the autonomy sports federations should be allowed when creating rules for a

<sup>59</sup> See n 27, above.

<sup>60</sup> See n 28, above.

<sup>61</sup> RTE previously paid €16 million for three years of rights; TV3 paid €6 million for just one package of rights. Samantha McCaughren, 'RTE poised to win battle for rights to GAA matches' *Sunday Business Post* (Dublin 10 February 2008) <http://archives.tcm.ie/businesspost/2008/02/10/story30359.asp> accessed 17 December 2009.

<sup>62</sup> *Deliège* (n 8) at [46].

<sup>63</sup> Weatherill (n 54).

particular sport.<sup>64</sup> Indeed, the *Oulmers*<sup>65</sup> case is an example of this type of approach in attacking sporting rules. At the time of writing, a similar case is mooted by Arsenal, which reportedly plans on suing the Netherlands football association for an injury suffered by Robin van Persie whilst on international duty with Holland.<sup>66</sup> One would expect Arsenal to raise the possibility of mandatory player release for championship qualifiers, friendly games and continental championships outside Europe without compensation for the employing club being a breach of Article 101 and/or 102 TFEU. These rules will now be assessed under the *Wouters* criteria as a result of the *Meca-Medina* judgment. Account will be taken of the overall context in which the rules were adopted and the objectives of the rules. The next question will be whether the restrictive effects are inherent in the pursuit of the objectives. Finally, it must be considered whether the restrictive effects are proportionate to the objectives.

Szysczak points out that the ‘pyramid structure’ of sports organisation, whereby one federation is placed in a de facto monopoly dominant position in the governance of the sport may be incompatible with EU competition law, and that *Meca-Medina* opens the door for the internal grievance mechanisms and disciplinary procedures created by such federations to be challenged under competition law.<sup>67</sup> The penalties imposed by such procedures must be proportionate to the objectives that the federation is seeking to achieve via the disciplinary procedures, and the process involved in grievance mechanisms must also be fair and proportionate. One could think of examples where a disciplinary rule has been applied in an inconsistent manner to different athletes, e.g. in professional rugby where suspensions for eye-gouging appear to differ between players, and often bear little relation to the nature of the offence itself. A long suspension for such an offence undoubtedly has an economic impact on the player concerned and his employing club and/or national rugby union.

In this author’s view, the ECJ’s judgment in *Meca-Medina* is a welcome clarification that most high-level sport is an economic activity and its organisation must comply with competition law. It is a practical recognition of the fallacy of demarking ‘purely sporting rules’ from rules having economic effects, in sports where the whims of monopoly sporting federations can lead to large shifts in revenue. The ECJ’s analysis was criticised by Subiotto, who opines that EU competition law should not have been applied to the IOC’s application of anti-doping rules, as the IOC was not acting as an ‘undertaking’ in carrying on the activity, as is required for the application of competition law to an entity. He claims that the IOC was not carrying on ‘economic activity’ in carrying out this function.<sup>68</sup> A rebuttal of this point would be to highlight that the IOC’s economic interest in clean sport is protected by applying stringent anti-doping rules. The perception of clean sport results in financial benefits for the IOC in terms of television money and sponsorship revenue. A cogent description of the economic consequences of anti-doping regulations is provided by Gregory,<sup>69</sup> when he points out that the cost to Manchester United of continuing to pay Rio Ferdinand’s wages during his eight month ban for failing to submit to a drugs test was £2.4

<sup>64</sup> Szysczak (n 21), p 105.

<sup>65</sup> *Oulmers* (n 2).

<sup>66</sup> Amy Lawrence, ‘Arsène Wenger’s legal action set to change the face of football’, *Guardian* (London 12 December 2009) <http://www.guardian.co.uk/football/2009/dec/12/arsenal-holland-legal-action-van-persie> accessed 17 December 2009.

<sup>67</sup> Szysczak (n 21), p 107.

<sup>68</sup> Romano Subiotto, ‘How a lack of analytical rigour has resulted in an overbroad application of EC competition law in the sports sector’ (2009) 2 ISLR 21.

<sup>69</sup> Gregory (n 53), p 51.



million. The additional financial cost to the club of his absence in terms of success on the field would be difficult to quantify. When Adrian Mutu failed a drugs test, Chelsea terminated his presumably lucrative contract and pursued him for damages. It would be extremely difficult to argue that anti-doping rules are purely sporting and not of an economic nature.

### Commission White Paper

Sports federations may no longer be able to point to a rule as ‘purely sporting’ in order to escape analysis under the Treaty, but as long as they can show that the rule satisfies the *Wouters/Meca-Medina* criteria, it will not breach Article 101(1) TFEU. *Meca-Medina* has informed the Commission’s 2007 White Paper on Sport,<sup>70</sup> which is a ‘road map’ for future EU action in sport. The Commission Staff Working Document<sup>71</sup> provides important background to the White Paper and gives an insight into the Commission’s thinking on the application of EU competition law to sport post-*Meca-Medina*. The Working Document is helpful in that it indicates the types of sporting rules and regulations which the Commission views as being in conformity with EU law, and also the approach it will take in the future in applying competition law to such rules.

The Working Document acknowledges the rejection by the ECJ of the argument that rules of a purely sporting nature fall outside the scope of EU competition rules, and also that all organisational sporting rules that determine the conditions for professional athletes, teams or clubs to engage in sporting activity as an economic activity are subject to the scrutiny of Articles 101 and 102 TFEU.<sup>72</sup> The specific features of sport will not be used to a priori exempt certain sporting rules from the application of the Treaty. The Working Document states that the specificity of sport is a factor which will be taken into account when applying the methodological framework set out in *Meca-Medina* to the sporting rule.<sup>73</sup>

The Commission helpfully sets out the methodology for applying Articles 81 and 82 EC (now Articles 101 and 102 TFEU) to rules adopted by sporting associations following *Meca-Medina*:

Step 1. Is the sports association that adopted the rule to be considered an “undertaking” or an “association of undertakings”?

a. The sports association is an “undertaking” to the extent it carries out an “economic activity” itself (e.g., the selling of broadcasting rights).

b. The sports association is an “association of undertakings” if its members carry out an economic activity. In this respect, the question will become relevant to what extent the sport in which the members (usually clubs/teams or athletes) are active can be considered an economic activity and to what extent the members exercise economic activity. In the absence of “economic activity”, Articles 81 and 82 EC do not apply.

Step 2. Does the rule in question restrict competition within the meaning of Article

<sup>70</sup> See n 13 above.

<sup>71</sup> See n 25 above.

<sup>72</sup> Working Document, p 36.

<sup>73</sup> *ibid*, pp 37-38.



81(1) EC or constitute an abuse of a dominant position under Article 82 EC? This will depend, in application of the principles established in the *Wouters* judgment, on the following factors:

- a. the overall context in which the rule was adopted or produces its effects, and its objectives;
- b. whether the restrictions caused by the rule are inherent in the pursuit of the objectives; and
- c. whether the rule is proportionate in light of the objective pursued.

Step 3. Is trade between Member States affected?

Step 4. Does the rule fulfil the conditions of Article 81(3) EC?<sup>74</sup>

This outline is useful in terms of legal certainty for sporting federations and their lawyers. We now know the approach that the Commission, and probably the EU courts, will take in this area in the future.

Interestingly, the Working Document contains a summary of the previous case-law of the courts and decision making of the Commission as regards rules made by sporting associations, and uses it to speculate on examples of sporting rules unlikely and likely to infringe Articles 81(1) and 82 EC.<sup>75</sup> Rules deemed unlikely to infringe the competition rules include rules concerning selection criteria for sport competitions (as in *Deliège*, where such rules did not restrict the athletes' freedom to provide services), 'at home and away from home' rules (as in *Mouscron*), rules concerning the multiple ownership of sports clubs (as in *ENIC*), rules concerning the composition of national teams (e.g. *Walrave*), anti-doping rules (as in *Meca-Medina*), rules regarding transfer windows (ruled on in *Lehtonen*) and also licensing systems governing participation in professional leagues. The Working Document sets out rules likely to infringe the Treaty competition provisions as: rules protecting sports associations from competition (as in the *FIA* case<sup>76</sup>), rules excluding legal challenges of decisions by sports associations before national courts, rules concerning nationality clauses for sports clubs (e.g. in *Bosman*), certain rules regulating the transfer of athletes between clubs (this was an issue in *Bosman* and also in the ensuing negotiations between the Commission and FIFA regarding FIFA's transfer rules), and rules regulating professions ancillary to sport (e.g. in *Piau*, where the CFI questioned FIFA's right to regulate players' agents, but acknowledged that in the absence of national law or self-regulation amongst agents the profession required regulation by some entity).

It must be noted that the Commissions assessment of the rules following *Meca-Medina* is consistent with the outcomes which arose from the historical case-law and decisions prior to that case. The rules which, pre-*Meca-Medina*, were declared by the Courts and

<sup>74</sup> *ibid.*, p 38.

<sup>75</sup> *ibid.*, Annex 1, pp 70-78.

<sup>76</sup> A complaint was made to the Commission that the Formula One governing body, the FIA, which had a regulatory monopoly in motor-racing activity, was violating the competition rules of the Treaty by preventing the emergence of a competitor motor-racing championship to the FIA's Formula One competition (see Commission press release IP/99/434 of 30 June 1999. The Commission reached a settlement with the FIA following the issuance of a Statement of Objections from the Commission (see Commission press release IP/01/1523 of 30 October 2001).

Commission to comply with the free movement and/or competition provisions of the Treaty are deemed by the Commission in the Working Document to comply with Articles 81 and 82 EC on application of the *Meca-Medina* criteria. This is notwithstanding the ECJ statement that a finding that a sporting rule does not violate Article 39 EC (on the free movement of workers) and/or Article 49 EC (on the freedom to provide services) does not exclude the rule from infringing Articles 81 and/or 82 EC.<sup>77</sup> Sporting rules assessed by the ECJ prior to *Meca-Medina* were ruled upon using free movement provisions of the treaty, not under the competition rules. One could be forgiven for asking whether the *Meca-Medina* judgment will have any practical impact on whether a sporting rule will infringe the competition rules, when the Commission seems to have adopted an 'as you were' position regarding sporting regulations and their compatibility with the Treaty.

## Conclusion

It is clear that there no longer exist rules of a 'purely sporting nature' which will automatically not fall under the scrutiny of EU competition law. Only trivial 'rules of the game' such as field dimensions or the duration of a game will not be considered and open to scrutiny by the Treaty. However, the Commission Staff Working Document pursuant to the White Paper on Sport would suggest that the legitimacy of existing sporting rules already decided upon by the Commission and the Courts will not change after being scrutinised under the *Meca-Medina* methodology. In the opinion of this author, the concept of rules of a 'purely sporting interest' is not quite dead following *Meca-Medina*. The specific nature of sporting rules, and the special treatment afforded to sports associations when making regulations, are still likely to be factors that the Commission will take into account when considering the legitimate objective pursued by the rule, whether any restrictions caused by the rule are inherent in the pursuit of that objective, and whether the rule is proportionate in relation to the objective.<sup>78</sup> While the ECJ's judgment in *Meca-Medina* is entirely realistic in terms of recognising the economic nature of most sporting activity, it remains to be seen whether the ECJ and the GC will continue to give sports federations broad autonomy in regulating sports, or whether Articles 101 and 102 will be applied stringently to sporting rules in the future. Unfortunately, the removal of the *Oulmers* case from the ECJ's pending case list means that the first opportunity for the court to apply *Meca-Medina* again has been lost. There is little doubt, however, that sporting rules are now ripe for further challenge under EU competition law, and further cases are inevitable.

<sup>77</sup> *Meca-Medina* at [32-33]. For further discussion on this point see Weatherill (n 53) at pp 649-650.

<sup>78</sup> The *Meca-Medina* criteria for assessing the compatibility of a sporting rule with the competition rules of the Treaty.



# THE RIGHT TO LIFE SHALL BE SECURED TO EVERYONE BY LAW: THE EXTENT TO WHICH THE EUROPEAN COURT OF HUMAN RIGHTS HAS DEVELOPED THE CONCEPT OF POSITIVE OBLIGATIONS IN RELATION TO ARTICLE 2

LOUISE MCCAMPHILL, LL.M.  
QUEEN'S UNIVERSITY BELFAST  
KMCCAMPHILL01@QUB.AC.UK

**ABSTRACT** Under the auspices of Article 2, the European Convention on Human Rights stipulates that “the right to life shall be secured to everyone by law.” Through analysis of case law, this paper investigates the extent to which the European Court of Human Rights has developed the concept of positive obligations in relation to Article 2. It is submitted that over time, the Court’s attitude has changed from one of apathy, towards an embracement of the concept; a claim corroborated by its innovative interpretation of the test laid down in *Osman v United Kingdom*. Although commonly criticised for being narrow in scope, it is my assertion that the Courts application of the test has been inventive and far reaching, extending the ambit of positive obligations under Article 2 to issues including disappearances and areas of environmental, custodial, individual and medical concern. Hence, it is argued that the court has created a framework which not only acts as a safety net, catching that which does not fall under Article 2’s negative side, but also provides a sound basis for further development of positive obligations in relation to Article 2; the ultimate result of which is, a more complete protection of the fundamental right to life.

Hailed as “one of the most fundamental provisions in the Convention”,<sup>1</sup> it is hardly surprising that an abundance of case law has been generated by Article 2 of the European Convention on Human Rights and Fundamental Freedoms. In order to ensure that “the right to life is secured to everyone by law,”<sup>2</sup> the European Court of Human Rights<sup>3</sup> has held that Article 2, which expressly imposes a negative obligation requiring authorities to refrain from unlawful killing, also imposes a positive obligation on States to protect life. As I endeavour to investigate the extent to which the Court has developed this concept I will illustrate the nature of positive obligations on a general level, consider their existence in the context of the Convention and focus on their development in relation to Article 2 through analysis of jurisprudence on the matter.

In a general context, an obligation is defined as “an act or course of action to which a person is morally or legally bound.”<sup>4</sup> Although suggesting that an obligation solely requires one to take a course of action, experience has shown that it may also require one to refrain from doing something. Consequently it is possible to identify an obligation as being positive, negative or a combination of both. In order to ensure that its requirements are met, it may be necessary to dismantle the obligation into a set of duties. These duties can exist in positive and/or negative form and it is through their examination that the scope of an obligation is revealed.

It has been noted that the European Convention of Human Rights is “mainly concerned with setting limits on the ability of State authorities to interfere with individual rights.”<sup>5</sup> This is demonstrated by the express words of the “rights and freedoms” enshrined in Section 1 of the Convention which impose mainly negative obligations and is representative of the classical view that leans towards imposition of negative rather than positive obligations.<sup>6</sup> Arguably the reason for this preference is because, contrary to positive obligations which impose, “actual expenditure and the deployment of resources,”<sup>7</sup> negative obligations do not require authorities to take action; thus imparting a lesser impact on allocation of resources. This, coupled with the fact that the Court has expressly declined to develop a general theory of positive obligations, has meant that the concept has largely been developed on a case by case basis.<sup>8</sup> Consequently it follows that almost all of the positive obligations in the Convention owe their existence to the judicial creativity of the Court.

Article 1, which states that, “High Contracting Parties shall secure to everyone within their jurisdiction the rights and freedoms defined within Section 1 of the Convention”<sup>9</sup> has generally been accepted as the source of legal authority permitting development of positive obligations in this way. Possessing an overarching status which on the one hand has a negative aspect, in that it requires States not to infringe Convention rights, and on the other relays positive duties to ensure that rights in the Convention are guaranteed,<sup>10</sup> Article

<sup>1</sup> *McCann and others v United Kingdom* (App no 18984/91) (1996) 21 EHRR 97, para [147].

<sup>2</sup> *European Convention on Human Rights and Fundamental Freedoms*, Article 2(1).

<sup>3</sup> Hereafter I shall refer to the European Court of Human Rights as the Court.

<sup>4</sup> Oxford Concise English Dictionary, (11th ed, revised) OUP, New York, (2008) p 986.

<sup>5</sup> Keir Starmer, “Positive Obligations Under the Convention” in Jeffery L Jowell and Jonathan Cooper (eds), *Understanding Human Rights Principles* (Oxford: Hart, 2001) p 139.

<sup>6</sup> Alastair Mowbray, *The Development of Positive Obligations under the European Convention on Human Rights by the European Court of Human Rights*, (Oxford: Hart, 2004) p 3 referring to Article 6(3) of the European Convention on Human Rights which requires States to provide free legal assistance to defendants in certain criminal cases, as an example of an “exceptional” right expressly placing a positive obligation on the State.

<sup>7</sup> *ibid*, p 4, citing Andrew Clapham, *Human Rights in the Private Sphere*, (Oxford, Clarendon Press, 1993) p 345.

<sup>8</sup> Starmer (n5) p139, see *Platform Artze für das Leben v Austria* (App no 10126/82) (1991) 13 EHRR 204.

<sup>9</sup> European Convention on Human Rights (n2).

<sup>10</sup> Clare Ovey and Robin C A White, *The European Convention on Human Rights* (4th ed, OUP, 2006) p 20.



1 can be described as the basic obligation that States enter into.

Its foundation rests on the principle of effectiveness, a “crucial jurisprudential tool”<sup>11</sup> that has arguably been a catalyst in the development of positive obligations, requiring a purposive interpretation of Convention articles in order to strive beyond achieving a merely theoretical protection – “The emphasis of the Convention is therefore on the “effective” protection of human rights, not the entrenchment of “theoretical” or “illusory” rights.”<sup>12</sup> Thus its recognition that, “a purely negative approach... cannot guarantee effective protection” arguably constitutes the theoretical basis for imposition of positive obligations in relation to certain Convention articles.<sup>13</sup>

Under the auspices of Article 1, and as one of the most fundamental rights in the Convention, the obligation to safeguard the right to life is a dual-edged sword. Its negative form is evident from the express words of Article 2, that require States to refrain from killing; the scope is obvious from the situations outlined in Article 2(2), in which it is deemed that death under these circumstances is lawful. Less obvious is the capacity of the positive obligation, although its existence can be identified from Article 2(1) which requires that “everyone’s right to life shall be protected by law.” Although it can be deduced that Article 2’s positive obligation has a protective function, the ambit of which has been determined by the Court in its adjudications, its development is dependent upon whether or not the “hurdle of effectiveness” is surpassed; essentially whether or not it is necessary to interpret Article 2 as imposing a positive obligation in order to achieve effective protection – “The object and purpose of the Convention as an instrument for the protection of individual human beings also requires that Article 2 be interpreted and applied so as to make its safeguards practical and effective.”<sup>14</sup> Therefore, in order to assess the extent to which the Court has developed positive obligations in relation to Article 2, it is necessary to conduct an analysis of the case law responsible for creating the duties falling under it.

The concept of positive obligations in connection with Article 2 was first recognised by the Court in *LCB v United Kingdom*.<sup>15</sup> In its assertion that, “the first sentence of Article 2(1) enjoins the State, not only to refrain from the intentional and unlawful taking of life, but also to take appropriate steps to safeguard the lives of those within its jurisdiction” the Court imposed an obligation on the State to take steps to protect the right to life where necessary.<sup>16</sup> This innovation constituted a major extension of Article 2’s ambit and created a basis on which many violations, previously only challenged via the negative route could now be challenged for failing to comply with the positive, hence providing for a more effective protection of the right to life.

That this case represented a move from an apathetic attitude towards positive obligations, to an espousal of the concept, is particularly potent when compared to the earlier decision of *McCann and others v United Kingdom*.<sup>17</sup> The Court had a choice between interpreting the failure to exercise strict control over security force operations involving use of lethal force,

<sup>11</sup> Mowbray, (n6) p 3 citing JG Merrills, *The Development of International Law by the European Court of Human Rights* (Manchester, MUP, 1993) p 102-3.

<sup>12</sup> Starmer (n5) p 145.

<sup>13</sup> *ibid*, p 139.

<sup>14</sup> *Özalp and others v Turkey*, (app no 32457/96) Judgement of 8th July 2004, para [32].

<sup>15</sup> *LCB v United Kingdom* (app no 23413/94) (1998) 27 EHRR 212.

<sup>16</sup> *ibid*, para [36].

<sup>17</sup> *McCann* (n 1).

as constituting a breach of its negative obligation, or as a breach of Article 2's positive obligation. In electing to follow the negative route it is plausible to contend that the Court demonstrated initial reluctance to develop Article 2's positive side and that in light of this cautious approach, *LCB v United Kingdom* can be viewed as a significant step towards its development.

The significance is also witnessed through a comparison of pre- and post-*LCB* case law. Contrary to the stance adopted in *McCann*, the Court's assertion in *Ergi v Turkey*<sup>18</sup> that responsibility of a State may be engaged where the authorities "fail to take all feasible precautions in the choice of means and methods of a security operation... with a view to avoiding, and... minimising incidental loss of civilian life" resembles their earlier promulgation in *LCB v United Kingdom*.<sup>19</sup> In *Ergi* it was held that the authorities had failed in their positive duty to protect the victim's life because of defects in the planning of the security forces operation; thus supporting the proposition that *McCann* could have been interpreted in the same way.

However, despite initiating a groundbreaking move in *LCB v United Kingdom* the Court did not grasp the opportunity to specify the "appropriate steps" to taken in order to safeguard life, indicating a preference that these be determined on a case-by-case basis. Although a breach was not found in this case due to the absence of evidence demonstrating a causal link between a father's exposure to radiation and his daughter's illness, analysis of the judgement indicates some of the duties required to fulfil the authorities' positive obligation.

It was the Court's suggestion that had this evidence existed, the authorities would have been under a duty to pass this information onto the applicant's parents, even if it would not have actually helped to reduce the risk to the applicant's life.<sup>20</sup> This daring proposition widens the scope of Article 2 and imposes a weighty burden on State authorities, requiring them to take reasonable measures to protect life even where it is known that the actions will be of little value. The merit in the submission found in the fact that it effectually forbids authorities from making value judgements, allowing the individuals concerned to do this for themselves.

It is notable that the Court did not make any such remarks in relation to the proposed duty to monitor the applicant's health. It is possible to deduce that where the causative link is established, State authorities are only under an obligation to monitor health where this will assist in reducing a risk to life. This narrower interpretation is arguably justified by resource considerations, the passage of information impacting less on resources than the monitoring of a person's health. Therefore it is possible to witness the limiting effect that such factors impose on the Court's development of Article 2's positive obligations. On balance, however, against a backdrop of indifference *LCB* established that positive obligations exist in relation to Article 2, thus widening the scope of protection afforded by it.

In *Osman v United Kingdom*<sup>21</sup> the Court held that the positive obligation to protect life extends beyond the minimum duty to put in place a legal framework providing effective

<sup>18</sup> Cited by Mowbray, (n 6), p 11 citing 1998-IV and C Buckley, "The European Convention on Human Rights and the Right to Life in Turkey" (2001) 1 Human Rights Law Review 35.

<sup>19</sup> *ibid*, p 12 citing European Commission Report of 20th May 1997, para [79].

<sup>20</sup> *LCB* (n 15), para [40], (*obiter dictum*).

<sup>21</sup> *Osman v United Kingdom* (app no 23452/94) (1998) 29 EHRR 245.

protection for Convention right<sup>22</sup> towards the inclusion of “a positive obligation on the authorities to take preventative operational measures to protect an individual whose life is at risk from the criminal acts of another individual.”<sup>23</sup> This is a significant step casting a wide net over situations that potentially constitute a breach of Article 2, and is hence a milestone in the Court’s development of positive obligations.

The Court, however, was anxious to impart that, “not every claimed risk to life can entail for the authorities a Convention requirement to take operational measures to prevent that risk from materialising.”<sup>24</sup> Consequently they devised a threshold to be met in order for it to be held that the authorities had failed in their duty to implement sufficient operational measures. It must be shown that they,

knew, or ought to have known at the time, of a real and immediate risk to the life of an identified individual or individuals from the criminal acts of a third party and that they failed to take measures within the scope of their powers which, judged reasonably, might have been expected to avoid that risk.<sup>25</sup>

This test attempts to define the ambit of protection afforded by Article 2's positive obligation to take operational measures and has been criticised as being very narrow in scope.<sup>26</sup> Upon closer inspection, however, it is arguable that the test is theoretically extensive in nature, as even if the authorities are not aware of the risk to an individual’s life they may still be held liable for a breach if it can be shown that they ought to have been aware of it. Thus it rejects the possibility of failure to find a breach based on subjective reasoning as it does not allow authorities to escape responsibility by using the excuse that they were unaware of such a risk.

The view that the test is of limited scope is buttressed by the requirement that the risk must be both “real and immediate.” On a strict application it would seem that even though the authorities may have known, or ought to have known of a risk, and had it within their power to take reasonable preventative action, they will not be deemed to have breached Article 2 if the risk was classified as being real but not immediate. Therefore the incorporation of the word ‘immediate’ arguably reduces the obligation’s remit. Perhaps the rationale for its inclusion was to avoid opening the floodgates of litigation. The Court’s warning that not every claimed risk to life falls within the duty to take preventative operational measures conceivable, however, would have sufficed and thus have eliminated the need to restrict the obligation’s ambit.<sup>27</sup>

Nevertheless, in light of the United Kingdom’s contention that the authorities’ failure to take preventative action should amount to “gross dereliction or wilful disregard of their duty to protect life” and their proposition that the obligation only arises where there “is a known risk of a real, direct and immediate threat to that individual’s life and where the authorities have assumed responsibility for his or her safety,”<sup>28</sup> the *Osman* Test may be described as theoretically broad. In their rejection of this proposal the Court expressed the

<sup>22</sup> For a discussion of this refer to Starmer, (n 5), p 147 and see *Dodov v Bulgaria*, (app no 59548/00, Judgement of 17th April, 2008) in which the Court held that there was a violation of Article 2 due to failure to provide effective judicial remedies with which to establish the circumstances of a disappearance and to hold to account, those responsible for endangering of life; basically a failure to put in place an effective legal framework.

<sup>23</sup> *Osman* (n 21) para [115].

<sup>24</sup> *ibid*, para [116].

<sup>25</sup> *ibid*, para [116]. Hereafter I shall refer to this as the *Osman* Test.

<sup>26</sup> Ovey and White, (n 10) p 63.

<sup>27</sup> *Osman* (n21) para [116].

<sup>28</sup> *ibid*, para [107].

opinion that:

... such a rigid standard must be considered to be incompatible with the requirements of Article 1 of the Convention... to ensure the practical and effective protection of the rights and freedoms laid down therein, including Article 2.<sup>29</sup>

Thus paying deference to Article 2's fundamental nature through the imposition of a lesser hurdle and the coverage of a more expansive area of conduct than the United Kingdom's proposition would ultimately encompass. In comparison it is evident that the two tests encapsulate opposite ends of the obligation's spectrum; their differences create a large gap between what is a very narrow and stringent test on the one hand, and what is a more open and wide one on the other. Thus, this corroborates the view that the *Osman* Test is a broad framework through which positive obligations may be further developed in relation to Article 2.

In light of this it is possible to reason that criticism of the Court's development of positive obligations should be directed at its application of the test to the facts of the case, rather than at the test itself. This view is arguably corroborated by the judgement in *Osman*, a case in which the applicants asserted that their family had become the target of a dangerous stalker who had broken into their home and shot and killed their husband and father.<sup>30</sup> They argued that "by failing to take adequate and appropriate steps" to protect them from the "real and known danger that Paget-Lewis posed, the authorities had failed to comply with their positive obligation under Article 2 of the Convention."<sup>31</sup> In support of this allegation the applicants pointed to a number of "warning signals" to bolster their claim that the police ought to have been aware of the risk and it was their contention that through contact with the headmaster of the school in which Mr Lewis taught, the police were aware that he was an "unbalanced, obsessive and aggressive individual who had stalked Ahmet Osman, taken photographs of him, plied him with gifts and even assumed his name."<sup>32</sup>

The applicants also sought to prove the authorities' failure to take reasonable measures to apprehend this risk by highlighting that the police had omitted to keep records of attacks on the Osman household, visits to the school and indications that the authorities were aware of Paget-Lewis' intention to commit murder.<sup>33</sup> Ultimately it was their contention that this casual and careless approach amounted to an "obvious inadequacy of the police response over a period of fourteen months and must be considered to amount to a grave dereliction of the authorities duty to protect life."<sup>34</sup>

Whilst the Court confirmed the applicants' contention that the "police were informed of all relevant connected matters" regarding Paget-Lewis' strange behaviour, they were of the opinion that this did not constitute behaviour alluding to a "real and immediate risk."<sup>35</sup> It was noted that although Paget Lewis' attachment to Ahmet was "most reprehensible from a professional point of view, there was never any suggestion that Ahmet Osman was at risk sexually from him, less so that his life was in danger."<sup>36</sup> They held that:

<sup>29</sup> *ibid.*, para [116].

<sup>30</sup> Ovey and White (n 10) p 63.

<sup>31</sup> *Osman* (n21) para [101].

<sup>32</sup> *ibid.*, para [103].

<sup>33</sup> *ibid.*, para [105].

<sup>34</sup> *ibid.*, para [103].

<sup>35</sup> *ibid.*, para [117].

<sup>36</sup> *ibid.*

... if it appeared to a professional psychiatrist that he did not display any signs of mental illness or a propensity to violence it would be unreasonable to have expected the police to have construed the actions of Paget-Lewis ... as those of a mentally disturbed and highly dangerous individual.<sup>37</sup>

In essence the Court concluded that the applicants had

... failed to point to any decisive stage in the sequence of events leading up to the tragic shooting when it could be said that the police knew or ought to have known that the lives of the Osman family were at a real and immediate risk.<sup>38</sup>

It appears that the decision was heavily influenced by the need to balance other Convention rights into the equation – “the police must discharge their duties in a manner which is compatible with the rights and freedoms of individuals,” and the need to uphold the presumption of innocence:

[T]hey cannot be criticised for attaching weight to the presumption of innocence or failing to use powers of arrest, search and seizure having regard to their reasonably held view that they lacked at relevant times the required standard of suspicion to use those powers or that any action taken would in fact have produced concrete results.<sup>39</sup>

Thus the influence of outside factors on the Court’s development of positive obligations in the context of Article 2 is clearly demonstrated.

It is surprising that the severe and unfortunate facts of this case did not result in the finding of a breach as, arguably it would have been possible for the authorities to have conducted a more thorough response to Mr Lewis’ bizarre behaviour without infringing on his presumed innocence or imposing “an impossible or disproportionate burden” on them.<sup>40</sup> Overall, it appears that the Court placed all their eggs in the psychiatrist’s basket, paying too little regard to the disturbing and odd behaviour of Mr Lewis which took place over a significant period of time. This observation is echoed by the dissenting minority who took the view that as Paget had already caused some harm and there was a strong indication that “more serious harm was to be foreseen.”<sup>41</sup> They advocated that the “professional experience that one is entitled to expect” required the police to exercise caution and take operational measures to protect those at risk and that in failing to do so they had breached Article 2.<sup>42</sup>

In light of this, it is possible that the conclusion forms the basis underlying the commonly held view that the *Osman* Test operates in too small a dimension to effectively provide protection of the right to life. The approach taken suggests that the Court was anxious to avoid an all-encompassing effect, preferring instead to control and manipulate the growth of positive obligations, leaving their development open to extension or reduction as future cases are decided. Overall, the analysis of *Osman v United Kingdom* illustrates that although the test itself is clear, its application causes reason for

<sup>37</sup> *ibid.*, para [118].

<sup>38</sup> *ibid.*, para [121].

<sup>39</sup> *ibid.*

<sup>40</sup> *ibid.*, para [116].

<sup>41</sup> *ibid.*, Opinion of Judge De Meyer.

<sup>42</sup> *ibid.*



interpretative debate, and the outcome suggesting that dramatic circumstances need to exist for the threshold to be crossed.<sup>43</sup>

An example of the extreme circumstances required is illustrated by reference to the cases of *Kaya v Turkey*<sup>44</sup> and *Kiliç v Turkey*<sup>45</sup> in which the Court held that the authorities had failed in their positive obligation to protect the life of the victims, a doctor and journalist respectively. In each case the applicants argued that as those assisting the PKK were regarded with suspicion by the authorities, the victims' links with the group demonstrated a risk to their lives. The Court accepted that as a journalist for Özgür Gündem, Kiliç was at serious risk of being targeted<sup>46</sup> – “the authorities were aware that those involved in the publication and distribution of Özgür Gündem feared that they were falling victim to a concerted campaign tolerated, if not approved by State officials.”<sup>47</sup> In *Kaya* a “real and immediate risk” was held to be demonstrated by strong inferences that the perpetrators were known to the authorities, through the witnessing of a suspected terrorist receiving assistance from the gendarmes and by the victim’s transportation through a series of official checkpoints.<sup>48</sup>

More significantly the Court conceded that these contentions were buttressed by a number of prevalent characteristics in Turkish society at the time, including the existence of a report supporting allegations that unlawful attacks were being carried out with the knowledge of, and in some instances, by the authorities. It was held that defects in investigations into these unlawful killings “undermined the effectiveness of the protection afforded by the criminal law,”<sup>49</sup> consequently emphasising the existence of a real and immediate risk and breaching the duty to take effective operational measures to protect the right to life.

The Court’s emphasis on the political climate of Turkey at the time demonstrates that Article 2’s reach extends beyond a mere consideration of the specific facts of each case to the inclusion of its context. It is the existence of this backdrop that distinguishes the cases of *Kaya* and *Kiliç* from the unfortunate event unattributable to any form of general practice, as witnessed in *Osman*. Consequently, it is plausible to believe that the Court is more willing to find a breach where the situation is indicative of a systematic practice of human rights abuse. The rationale was possibly one based on the view that this is an indirect way to curb wide-scale breaches of Article 2 and thus offer extended protection. In doing so the Court has potentially developed a special category of cases occupying a privileged place under the umbrella of Article 2.

In its statement that “it is incumbent on the State to account for any injuries suffered in custody, which obligation is particularly stringent where that individual dies”<sup>50</sup> the Court has placed those in custody in a similar taxonomy. *Edwards v United Kingdom*,<sup>51</sup> which concerned the murder of an inmate by a dangerous and mentally ill prisoner placed in his cell is demonstrative of the Court’s effort to develop positive obligations in relation to

<sup>43</sup> cf *Branko Tomasic and others v Croatia*, (app no 46598/06) Judgement of 15th January 2009, a case similar in facts to *Osman* in which a man murdered his partner and child having previously made death threats. It was held that these were extreme circumstances and that as the authorities response was inadequate, they had failed to fulfil their positive obligation to protect the right to life under Article 2.

<sup>44</sup> *Mahmut Kaya v Turkey*, (app no 22535/93) Judgement of 28 March 2000.

<sup>45</sup> *Kiliç v Turkey* 33 EHRR 1357.

<sup>46</sup> *ibid*, para [57].

<sup>47</sup> *ibid*, para [66].

<sup>48</sup> *Kaya* (n 44) para [87].

<sup>49</sup> *ibid*, para [98].

<sup>50</sup> *Paul and Audrey Edwards v United Kingdom* (app no 4647/99) (2002) 35 EHRR 487 para [56].

<sup>51</sup> *ibid*.

Article 2. This case saw the Court depart from the requirement that there must be a “decisive stage in the sequence of events”<sup>52</sup> pointing to a real and immediate risk of which authorities were or ought to have been aware, to the interpretation of it as encompassing situations in which a range of events have accumulated to create a risk to life. In doing so the Court increased the number of situations in which authorities must take active steps to secure the right to life and thus extended the ambit of positive obligations in relation to Article 2.

A welcome side effect of this development was the confirmation that the existence of a breach was not solely assessed in terms of the action or inaction of the prison authorities but also on an investigation of the events leading up to the victim’s death. Thus the Court’s analysis included a consideration of the “failure of the registrar to consult Richard Linford’s notes... and the failure of the police, prosecution or Magistrates Court to take steps to inform the prison authorities... of Richard Linford’s suspected dangerousness and instability.”<sup>53</sup> It is clear that the Court viewed this “series of shortcomings in the transmission of information” as the cause of Edwards’ death because its successful passage would have illuminated the risk and placed the authorities in a position to take reasonable steps to protect Edwards.<sup>54</sup> Therefore, the Court reached the decision that, as the authorities ought to have been aware of the risk to the victim’s life, they ought to have taken reasonable steps to alleviate that risk and in their failure to do so had breached Article 2.

Thus, *Edwards v United Kingdom* constitutes a milestone in the Court’s development of positive obligations, not only confirming the nod given towards the importance of passing on information in *LCB v United Kingdom*, but also preventing concealment behind the veil of other’s failures. Therefore, the case is responsible for the easier establishment of a “real and immediate risk” through exploration of a range of circumstances, rather than by identification of one defining moment. Through this extended interpretation of the *Osman* Test, the Court has widened the outreach of positive obligations and thus increased protection of the right to life afforded by Article 2.

The duty to pass on information was regarded as integral to the authorities’ positive obligation to protect the life of an inmate in the case of *Tarariyeva v Russia*.<sup>55</sup> It was held that as a result of “inadequate and defective medical assistance”<sup>56</sup> on the part of both the prison and public hospital involved, the deceased had died whilst in custody. The Court identified a number of causal links<sup>57</sup> between Mr Tarariyeva’s death and the inadequate medical care which was administered, one being failure on the part of the medical authorities to keep a record of his condition and in turn pass on this vital information to those dealing with the patient – “the state hospital doctors provided their colleagues in the prison hospital with an incomplete medical record... information crucial for the proper assessment of the patient’s condition was missing.”<sup>58</sup> Consequently, as the deceased had a history of ill health of which the authorities were aware,<sup>59</sup> the Court deemed such shortcomings inexcusable and, as a result, a contributory factor in their failure to fulfil the positive obligation to protect the right to life of the deceased.

<sup>52</sup> *Osman* (n 21) para [121].

<sup>53</sup> *Edwards* (n50) para [61].

<sup>54</sup> *ibid.*

<sup>55</sup> *Tarariyeva v Russia*, (App no 4353/03), Judgement of 14th December 2006.

<sup>56</sup> *ibid.*, para [68].

<sup>57</sup> *ibid.*, para [87].

<sup>58</sup> *ibid.*, para [83].

<sup>59</sup> *ibid.*, para [88].

This case is also significant in that it demonstrates the Court's application of positive obligations in the context of provision of medical care. In holding that "positive obligations require States to make regulations compelling hospitals, whether private or public, to adopt appropriate measures for the protection of patients' lives" the Court significantly extended the ambit of Article 2.<sup>60</sup> The Court indicated that this positive obligation encompasses duties on health authorities to ensure appropriate follow up care for patients<sup>61</sup> and the provision of adequate medical facilities;<sup>62</sup> both of which were held to be conspicuously lacking in this case. Therefore, along with the fact that the prison hospital could not provide blood transfusion facilities, the decision to discharge the patient early was found to be the "main cause of aggravation of Mr Tarairyev's condition."<sup>63</sup> This is thus an innovative and measurable interpretation of the positive obligation to protect the right to life that demonstrates the permeation of areas concerning medical practice; a relatively "unexplored"<sup>64</sup> area which will no doubt be further developed in the future.

The protection afforded by Article 2 was further extended by the Court in *Keenan v United Kingdom*<sup>65</sup> and in *Uçar v Turkey*<sup>66</sup> when it was held that the duty to take reasonable measures to protect life includes situations in which the risk emanates from the individual themselves. In doing so the Court rejected the British Government's assertion that "the principles of dignity and autonomy should prohibit any oppressive removal of a person's freedom of choice and action" and advocated that whilst "restraints will inevitably be placed on the preventative measures taken by the authorities," general measures and precautions exist that will "diminish the opportunities for self harm, without infringing on personal autonomy."<sup>67</sup>

Despite breaking existing boundaries set by the Court's previous interpretations, this development should be viewed with a cautious optimism as it is impossible to prevent such a risk materialising in every context in which it occurs. Therefore, further development of this duty may solely extend to medical and social welfare situations where it is more likely that the authorities have, or ought to have, an awareness of the individual's mental disposition to commit suicide. Existence of this limit is supported by the Court's warning that not every claimed risk to life falls within the duty to take preventative operational measures<sup>68</sup> and corroborates classification of custodial cases in a special category to which stringent protection is afforded.

The Court's broad interpretation of the requirement that the risk be both "real and immediate," however, is unlikely to confine itself to such a narrow remit. Despite concluding that the risk the deceased posed to himself was not immediate, the Court reasoned that "the variations in his condition required that he be monitored carefully in the case of sudden deterioration,"<sup>69</sup> suggesting that an unpredictable risk will suffice. It is possible to envisage this interpretation translating easily into a wide range of situations, including that of *Osman v United Kingdom*,<sup>70</sup> where the unpredictable manner in which

<sup>60</sup> *ibid*, para [74].

<sup>61</sup> *ibid*, para [80].

<sup>62</sup> *ibid*, para [87].

<sup>63</sup> *ibid*, para [84].

<sup>64</sup> Karen Reid, *A Practitioner's Guide to the European Convention on Human Rights*, (2nd edition, Sweet and Maxwell, 2004) p 505.

<sup>65</sup> *Keenan v United Kingdom*, (App no 27229/95) 33 EHRR 903.

<sup>66</sup> *Uçar v Turkey*, (App no 52392/99) Judgement of 11th April, 2006.

<sup>67</sup> *Keenan* (n 65) para [92].

<sup>68</sup> *Osman* (n 21) para [116].

<sup>69</sup> *Keenan* (n 65) para [96].

<sup>70</sup> *Osman* (n 21).

Paget-Lewis acted could have been interpreted as posing a real and immediate risk to the lives of the victims. Therefore, through innovative interpretation, the Court potentially broadened the realm of situations caught by the test, thereby developing and extending the ambit of Article 2.

Another milestone in the Court's development of positive obligations concerns the application of the concept in the context of disappearances. In the cases of *Koku v Turkey*<sup>71</sup> and *Osmanoğlu v Turkey*<sup>72</sup> the Court was of the opinion that the abduction and disappearance of the victims was synonymous with a risk to their lives<sup>73</sup> and that as a result, the authorities failure to obtain more information from potential witnesses, to verify whether the abductors were in fact police officers, to alert the police and gendarmerie checkpoints and failure to conduct an inspection of their headquarters, constituted an inadequate response to this risk.<sup>74</sup> Thus it was held that the authorities had failed to protect the life of the disappeared and as a result had breached their positive obligation under Article 2.

This development was not without controversy. The minority in *Osmanoğlu* voiced the objection that a breach was found under the substantive element of Article 2, rather than under the existing procedural obligation to conduct an investigation into unlawful deaths. The basis of their argument was that the "respondent State should not be held responsible... if it is established that the disappearance did not occur when the victim is under the control of the authorities and when the State is not involved in the disappearance."<sup>75</sup> This position is irreconcilable with founding case law on the matter<sup>76</sup> and ridicules the test laid down in *Osman* in which such a stipulation was not impressed. If true, it would significantly narrow the ambit of positive obligations in relation to Article 2 and render the protective function as practically ineffective, thus ignoring Article 1's status as an overriding obligation and Article 2's fundamental importance.

The dissenters also sought to support their view with the argument that the positive obligation is preventative in nature and "relates to a phase before such an incident occurs."<sup>77</sup> The proposition that a positive obligation only exists prior to disappearance in order to ensure its prevention and that after this, it becomes a procedural duty outside the scope of the obligation to take measures to protect life, is ill advised. There is no reason to curtail the reach of positive obligations at this point as the risk to life does not come to an end once a person is abducted; to the contrary it is increased. This argument was confirmed by the Court in its statement that:

... the action which was expected from the domestic authorities was not to prevent the disappearance of the applicant's brother – which had already taken place – but to take preventive operational measures to protect his life which was at risk from the criminal acts of other individuals.<sup>78</sup>

<sup>71</sup> *Koku v Turkey* (App no 27305/95) Judgement of 31st May 2005.

<sup>72</sup> *Osmanoğlu v Turkey* (App no. 48804/99) Judgement of 24 January 2008.

<sup>73</sup> Cf *Gekhayevea v Russia*, (app no 1755/04) Judgement of 29th May 2008) which found a violation of Article 2 for failure to conduct adequate investigation into the disappearance of two sisters, whose lives were deemed to be at risk from the moment of their disappearance, further emphasised by the lapse of three years and five months. It was held that the positive obligation to investigate was not met because of delays, an adjournment and a failure to exercise adequate investigative powers. Hence demonstrating the gravity placed on positive obligations in the context of disappearances.

<sup>74</sup> *Osmanoğlu* (n 72) para [79-80].

<sup>75</sup> *ibid*, Joint Partly Dissenting Opinion of Judges Türmen, Vajic and Steiner.

<sup>76</sup> See *LCB* (n 15).

<sup>77</sup> *Osmanoğlu* (n 72) dissenting opinion.

<sup>78</sup> *Koku*, (n 71) para [132].

Thus a positive obligation to investigate disappearances exists and applies from the moment of a disappearance because it is from this moment that the risk to life exists.

In order to aid their argument the dissenters also pointed to the fact that “the Court has always interpreted this duty of the State rather narrowly” and that it does not “impose a requirement that a State must necessarily succeed in locating and prosecuting perpetrators of fatal attacks.”<sup>79</sup> In rebuttal of these points it should be noted that the Convention is viewed as a “living instrument” and that whilst the Court may have interpreted narrowly in the past, this is no reason to curtail future development.<sup>80</sup> It is evident that the dissenters have disregarded the Court’s warning that not every risk to life constitutes a breach and failed to comprehend the fact that “the obligation to take steps to protect the right to life is not an obligation of result but of means.”<sup>81</sup> Consequently the Court’s determination that it is sufficient for an applicant to demonstrate that the authorities did not do all that could reasonably be expected of them in order to avoid a real and immediate risk to life should allay any requirement of a mandatory solution leading to identification of those culpable.<sup>82</sup>

Although framed in the context of a duty to take operational measures, the positive obligation requiring an investigation of the circumstances surrounding disappearances is akin to that of investigating suspicious deaths. In choosing to encompass this duty as a positive obligation, however, the Court have matched the framework to the aim sought to be achieved; use of the procedural component which is concerned with providing a remedy for breach of Article 2 by way of fact-finding investigations into suspicious deaths is an inappropriate vehicle when compared to the substantive positive obligation which is protective and preventative in nature. Therefore, relegation of the duty to investigate disappearances as a procedural duty, rather than a substantive one, would contradict the fundamental importance of Article 2 and consequently the Court can be commended for developing positive obligations in a proactive manner.

The Court recognised this fundamental and all-encompassing nature of Article 2 in their acknowledgement of a link between human life and the environment and through establishment of a duty on public or private bodies to take reasonable operational measures to protect the right to life where it is at risk from inherently dangerous but lawful activity. Building on earlier jurisprudence,<sup>83</sup> the Court found in *Öneriyildiz v Turkey*<sup>84</sup> that there was a positive obligation on authorities to protect those who had died as a result of an explosion from a municipal tip which had engulfed much of the shanty town below it. It was found inexcusable that regulatory law requiring installation of a ventilation system had not been complied with, especially in light of available information that specifically referred to the imminent danger.<sup>85</sup> Thus it was held that the Turkish authorities had breached their positive duty under Article 2 to protect the right to life.

In reaching this decision the Court noted that, insertion of ducts would “have been a much better reflection of the humanitarian considerations”<sup>86</sup> and illuminated the importance in maintaining public confidence, adherence to the rule of law and the prevention of any apparent tolerance of collusion in unlawful acts.<sup>87</sup> This suggests that the

<sup>79</sup> *Osmanoğlu* (n 72) dissenting opinion.

<sup>80</sup> *Tyrer v United Kingdom* (App no 5856/72) 2 EHRR 1, para [31].

<sup>81</sup> *Osmanoğlu* (n 72) para [77].

<sup>82</sup> *ibid*, para [77].

<sup>83</sup> cf *Guerra v Italy* (app no 14967/89) (1998) 26 EHRR 357 as cited in *Ovey and White* (n 10) p 64.

<sup>84</sup> *Öneriyildiz v Turkey* (app no 48939/99) (2004) 41 EHRR 325.

<sup>85</sup> *ibid*, para 98 referring to para [13] of same.

<sup>86</sup> *ibid*, para [107].

<sup>87</sup> *ibid*, para [96].



fundamental nature of Article 2, whilst inherently important in substance, has wider implications; the importance in promoting a visual condemnation of unlawful acts. Therefore the case is significant, not only because it pays deference to the importance of protecting life, on both substantive and notional levels, but also because it recognises that, “protection of the environment must be considered as a vital aspect of the right to life as without a sound environment it would not be possible to sustain an acceptable quality of life or even life itself.”<sup>88</sup>

Perhaps the most revealing judgement of the Court is that of *Pretty v United Kingdom*<sup>89</sup> which demonstrates the Court’s consistent approach to the development of positive obligations in relation to Article 2. Within it we see that the obligation placed on the State to protect the right to life is of such gravity that the Court is unwilling to consent to the diametrically opposite right, the right to die. The rationale for such a position is two-fold. Firstly there is a practical basis, in that the right to life is unconcerned with the quality of life or what a person chooses to do with it so long as they are still alive, thus invoking a purposive interpretation of the right.<sup>90</sup> They also rely on a literal interpretation to corroborate their position, deeming the right to die as being untenable without a distortion of the language of Article 2, and thus it is beyond the Court to allow it.<sup>91</sup> Overall, the Court seems to take the view that the right to life is so fundamental and important that to allow for a right to die would be an unthinkable corruption of the values upheld by the Convention.

On an overall analysis, the European Court of Human Rights has been instrumental in fostering recognition of Article 2’s fundamental importance through its establishment of positive obligations. Analysis of jurisprudence indicates that the Court’s interpretation of the *Osman* Test has followed a trend conducive to their effective development; this is demonstrated by a move from a restricted construction to a broad one permeating areas of environmental, custodial, individual and more recently, medical concern. A consequence of this extended interpretation is the creation of a subsidiary role for positive obligations, that of a safety net catching that which does not fall under Article 2’s negative side. Thus, this has the effect of invoking the protection of Article 2 in a wider sphere.

In conclusion, through its innovative interpretation of Article 2, the Court has not only acknowledged the cardinal importance of the right to life and extended the ambit of positive obligations accordingly, but has also created the potential for their future development through the framework of the *Osman* Test. Thus the concept is a creature of jurisprudential evolution that in light of Article 2’s fundamental importance will most likely continue to grow and permeate a wide range of areas, especially the relatively virgin area of medical services – “This judicial creativity is a worthy reflection of the right to life and there is considerable potential for future development in the emerging obligation on states to provide medical services.” Thus, it can be concluded that in its development of positive obligations under Article 2, the European Court of Human Rights has rigorously promoted the ideal that “the right to life shall be secured to everyone by law.”

<sup>88</sup> Jérémie Gilbert, “Environmental Degradation as a Threat to Life: A Question of Justice?” (2003) 6 Trinity CL Rev 81-97, p[83].

<sup>89</sup> *Pretty v United Kingdom* (App no 2346/02) (2002) 35 EHRR 1.

<sup>90</sup> *ibid.*, para [39].

<sup>91</sup> *ibid.*

<sup>92</sup> Mowbray, (n 6) p 41.

