

Aberdeen Student Law Review

**With thanks to our sponsors
Stronachs LLP**

**September 2013
Volume 4**

www.abdn.ac.uk/law/aslr

THE EDITORIAL BOARD

2012 - 2013

Managing Editor

Ilona Cairns

Assistant Editors

Jayne Holliday

Stephen Ooi

Editors

Campbell Stuart

Lillie Fraser

Felix Okpe

Catherine Guthrie

Tom Croy

Graham MacDonald

Luke Burgess-Shannon

FOREWORD BY THE HON. LORD WOOLMAN

SENATOR OF THE COLLEGE OF JUSTICE

Every legal system needs constant appraisal. Legislators, judges and practitioners are all involved in that task. But academic comment is also vital. Legal authors have the opportunity to stand back and cast a critical eye on matters. They may analyse a particular decision. Equally, they may assess a developing trend in a particular area of the law. Such scholarship is extremely valuable to Scots Law.

The *Aberdeen Student Law Review* is now building up a significant catalogue of articles. A glance at the contents page of the present volume shows the wide range of contributions. The authors have addressed topics in public law and private law; European law and Scots law; substantive law and evidence. It is clear that scholarship is thriving in Old Aberdeen and that the law school is in good heart.

Stephen Woolman

August 2013

INTRODUCTION TO VOLUME FOUR

The Aberdeen Student Law Review (ASLR) was established in 2010 to showcase the work of students and alumni of the Law School at the University of Aberdeen. The fact that the journal is now in its fourth year of publication is testament to the dedication and exceptional ability of these students, both past and present. It is also a reflection of the high standard and breadth of legal education that students receive at this University. As with past years, we have endeavoured to include articles that deal with diverse and topical legal issues and that are appealing and relevant on both a domestic and international level. We believe this is in keeping with the Law School's aspirations with respect to legal scholarship.

Ensuring diversity and balance between articles was stimulating but was not without its challenges. On a few occasions our commitment to variety and topicality meant that submissions of a high standard were not accepted for publication. We are genuinely appreciative of all the students who submitted their work to the ASLR. Of course, we are especially thankful to the authors of the articles included in this issue for working closely and diligently with the editorial board to bring their article up to the highest possible standard. Aside from the authors, there are many others who made the process of compiling this issue both enjoyable and rewarding. Enormous thanks go to Colin Mackie and Caroline Hood who offered considerable support and practical assistance throughout the year, and to Leanne Bain and Dominic Scullion for their helpful insights and continued interest in the ASLR. Peer reviewers also play a key role in the success of this journal and we are grateful to everyone who took the time and effort to review submissions. Thanks must also be extended to numerous staff members in the Law School, most notably Anne-Michelle Slater and Sarah Duncan, who recognise the enormous value in this publication and are always willing to lend a hand.

Last but certainly not least we are indebted to our sponsors, Stronachs LLP, whose encouragement and generosity is pivotal to the reputation and academic strength of the ASLR. With their support, we warmly welcome you to the fourth edition.

The Editorial Board

August 2013

Contents

Articles

Proceed with Caution(s): A Critique of the Carloway Review's Rejection of Statutory Adverse Inference Provisions in Scottish Criminal Law	Philip Glover	1
---	---------------	---

The Roma Minority & Free Election Rights in Bosnia and Herzegovina: The Problem of Constitutional Reform	Jennifer Paton	23
--	----------------	----

The Utility of the NESS Test of Factual Causation in Scots Law	Euan West	39
--	-----------	----

Analysis

Execution in Counterpart in Scots Law	Caroline Hood	63
---------------------------------------	---------------	----

Pragmatism in the European Union: Recasting the Brussels I Regulation to Ensure the Effectiveness of Exclusive Choice-of-Court Agreements	Justin Cook	76
---	-------------	----

Gender Mainstreaming & 'Equality Proofing' in British Law-Making: A Comment on the Impact of the Equality Act 2010	Ilona Cairns	92
--	--------------	----

Case Comment

<i>Royal Bank of Scotland v Stuart Hill</i>	Katherine Anderson	105
---	--------------------	-----

Book Review

Child Abduction within the European Union	Jayne Holliday	114
---	----------------	-----

News from Aberdeen Law School

The Aberdeen Law Project AGM 2013	Malcolm Combe	117
-----------------------------------	---------------	-----

Guidelines for Contributors		121
------------------------------------	--	-----

Proceed with Caution(s): A Critique of the Carloway Review's Rejection of Statutory Adverse Inference Provisions in Scottish Criminal Law

PHILIP B. GLOVER*

Abstract

Scottish criminal procedure (unlike the rest of the United Kingdom and the Republic of Ireland) currently contains no statutory provisions permitting courts to draw adverse inferences from an accused's silence during police questioning or at their subsequent trial. The law in this area was recently examined in the Carloway Review, which rejected the introduction of statutory adverse provisions on two principal grounds, both of which are analysed in depth and contested in this article. This article contends that the Review Team's reasoning was problematic, insofar as it ran contrary to the opinions of many Consultation respondents and employed analysis and final decision-making based on selective quotation from leading academic analysis on the subject. It is suggested that Scotland can and should enact adverse inference provisions using an amended system of police cautions that are compliant with the European Convention on Human Rights ('ECHR'). Carefully constructed cautions, in conjunction with the new statutory disclosure and legal assistance regimes now in place in Scotland, would ensure that an accused person's vulnerability would not be unduly compromised.

1. Introduction

Over eighteen months have elapsed since the publication of the Carloway Review's Report and Recommendations.¹ The Review represented the Scottish government's principal response to the controversy arising out of the UK Supreme Court's decision in *Cadder*.² Prior to publication, the Review team had circulated a Consultation Document,³ before considering fifty one responses.⁴ As regards the idea of

* The author is a 2012 LLB Honours Graduate of the University of Aberdeen, and a PhD candidate researching the utility of the Regulation of Investigatory Powers Act 2000 and the Regulation of Investigatory Powers (Scotland) Act 2000. I am deeply indebted to the support of Professor Peter Duff at the University, both in guiding my research and critically reviewing this article and the dissertation it is edited from.

¹ The Carloway Review, *Report and Recommendations*, 17 November 2011.

<<http://www.scotland.gov.uk/Resource/Doc/925/0122808.pdf>> accessed 10 October 2012.

² *Cadder v HM Advocate* [2011] SC UKSC 43. The controversy is not revisited here.

³ The Carloway Review, *Carloway Review Consultation Document*, 8 April 2011

<<http://www.scotland.gov.uk/Resource/Doc/925/0116090.pdf>> accessed 7 October 2012.

introducing adverse inference provisions ('AIPs') in statutory form in Scotland, their final recommendation was that 'no change is made to the current law of evidence that prevents inferences being drawn at trial from an accused's failure to answer questions during the police investigation.'⁵ The Review provided two principal reasons. The first was that '(...) the introduction of adverse inference would not fit well with the presumption of innocence, the right to silence and the privilege against self-incrimination as understood and applied in Scotland'⁶ (an essentially conceptual argument). The second was a more practical argument, namely that '[i]nstead of promoting efficiency and effectiveness, it would bring unnecessary complexity to the criminal justice system.'⁷

This article, in disagreeing with the Review's reasoning, will argue that Scottish criminal law should, principally through the introduction of appropriately worded European Convention-compliant caution(s), be adjusted to permit the drawing of adverse inferences from an accused's silence in circumstances where this complements other proven incriminatory evidence. This is not an isolated or new position; it is supported by several Consultation respondents.⁸ This article seeks to refocus critical analysis on the factual realities and utility of statutory AIPs, and in doing so lends weight to the view expressed by the Scottish Law Commission ('SLC'), one of the primary respondents to the Consultation, that the law regarding AIPs in Scotland should be re-examined.⁹ Prior to setting out the case for AIPs, this article initially agrees with the prescient observations made by Raitt,¹⁰ and will attempt to reinvigorate a '(...) reflective argument vital to the law reform process'¹¹ that she and other concerned consultation respondents correctly forecast would be excluded.¹² It also sympathetically acknowledges that a principal unpublished reason for rejecting statutory AIPs may have been the lack of available time to properly

⁴ The Carloway Review, *Responses to the Consultation Document*, 11 August 2011 <<http://www.scotland.gov.uk/About/CarlowayReview/Consultationprocess>> accessed 7 October 2012.

⁵ The Carloway Review, *Report and Recommendations* (n 1) 38.

⁶ *ibid* [7.5.26].

⁷ The Carloway Review, *Report and Recommendations* (n 1) [7.5.26].

⁸ The Carloway Review, *Responses to the Consultation Document* (n 4). See for example Raitt, at 61, the Association of Police Superintendents, at 102, and perhaps most importantly, the SLC at 200. The joint submission of Leverick & Farmer, at 34, states (from a stance opposing AIPs): '(...) if they are to be permitted at all, adverse inferences should only be permitted where there has been disclosure of the police evidence against the suspect.' This author agrees.

⁹ The Carloway Review, *Responses to the Consultation Document* (n 4) 202.

¹⁰ FE Raitt, 'The Carloway Review: An opportunity lost' (2011) 15(3) *Edinburgh Law Review* 427. At 429, she states: '(...) there must be real doubts about whether [Lord Carloway's agenda] can be achieved. The Review timetable permitted a consultation period of eight weeks. It is difficult to see how complex matters such as corroboration and the inference from silence can be fully explored in the proposed condensed schedule.'

¹¹ *ibid*.

¹² *ibid*. It is stressed that the degree of expertise and authority of those in the Review team is not in doubt. However the breadth of their remit and the timescale for completion led to criticism which was on occasion damning. See for example the Carloway Review, *Responses to the Consultation Document* (n 4) 319 for the succinct response of the Faculty of Advocates, who politely declined to respond. See also the articulate, measured criticisms of the Senators of the College of Justice, at 359.

consider the merits of a change in law in sufficient depth.¹³ In building a case for statutory AIPs in Scotland, it is conceded that the incidence of silence at police stations is statistically low¹⁴ and that commentators have drawn attention to the marginal ‘utility...and justification’¹⁵ of AIPs in other jurisdictions. However it is submitted here that the same could be said of the number of petitions to the *nobile officium* in Scotland. The fact that a tool for securing justice is only used infrequently does not mean that it should not remain available for use in appropriate circumstances.

In constructing my argument, the Carloway Review’s published research summarising the current Scottish position will be critically analysed, with reference to both case law and leading texts. It will be shown that there is a degree of AIP currently present in Scotland, but that this practice appears both inconsistent and under-utilised. This incoherency appears partially attributable to an unwritten ‘doctrine of fair play’ and sometimes conflicting judicial precedent that predates (and therefore fails to embrace) the ECHR and its jurisprudence tolerating AIPs (with safeguards) within an overall Article 6 right to fair trial proceedings.¹⁶ It will be suggested that these factors have led to a culture of resistance to the drawing of adverse inferences even though no actual prohibition exists. Next, in order to highlight Scotland’s unique position within the UK, a brief outline of the codified provisions existing in the neighbouring jurisdictions of Northern Ireland, England¹⁷ and the Republic of Ireland will be provided.¹⁸ The inherent *discretionary* nature of these AIPs will be highlighted in order to refocus analysis on the fact that such AIPs are discretionary, rather than *rule-based*, as is often suggested by opponents. It will be shown that this fundamental misconception as to how AIPs are applied has permeated the arguments of opponents to the extent that even the question relating to AIPs in the Consultation document was loaded in favour of those supporting the *status quo*.

As several opposing Consultation responses selectively cite Roberts and Zuckerman¹⁹ to support their positions, the actual stated position of these authors regarding AIPs will be reaffirmed. This will reveal that not one of the selectively quoted rationales underpinning the arguments of opposing respondents actually stands up to detailed scrutiny. It will then be shown that the Review’s first reason for rejecting AIPs reflects elements of the potential rationales identified and discredited

¹³ As regards depth, comparison is invited between the background, history lesson, consideration and final conclusions on the corroboration requirement in Scots Law (46 pages) in the Carloway Review, *Report and Recommendations* (n 1) [7.1], with the 13 pages offered on Adverse Inference [7.5].

¹⁴ P Roberts & A Zuckerman, *Criminal Evidence*, (2nd edn, OUP 2009) 578.

¹⁵ I Dennis, ‘Silence in the police station: The marginalisation of section 34’ [2002] (Jan) Criminal Law Review 38.

¹⁶ Council of Europe, *European Convention on Human Rights as amended by Protocols 11 & 14*, supplemented by Protocols Nos. 1, 4, 6, 7, 12 and 13 <http://www.echr.coe.int/NR/rdonlyres/D5CC24A7-DC13-4318-B457-5C9014916D7A/0/Convention_ENG.pdf> accessed 13 March 2013.

¹⁷ References to ‘England’ and ‘English legislation’ include Wales.

¹⁸ The outline of these provisions will be merely descriptive. Space constraints preclude discussion of the provisions in detail. Further, as will be suggested later in this article, criticisms of such provisions are of limited relevance to the creation of new AIPs for Scotland.

¹⁹ Roberts & Zuckerman, *Criminal Evidence* (n 14) 545-580.

by Roberts and Zuckerman. The Review's reasoning will be examined and rebutted, with additional scrutiny of the strong oppositional view expressed by Ferguson,²⁰ who appears to have had a significant influence on the Review's overall conclusions. The seminal European Court decision in *Murray*²¹ will be revisited to bolster my proposition that, as AIPs have been sanctioned by Europe's highest court (who have previously dictated adjustments to Scottish criminal law), there is little justification for not introducing AIPs in Scotland as an additional effective tool for securing justice in circumstances deemed appropriate. The subjective nature of complexity, coupled with the limited relevance of alluding to the English AIPs model when opposing AIPs for Scotland, will be demonstrated. An implicit oppositional rationale underpinning the Review's 'complexity' argument will also be discussed.

Finally, a personal view as to how a system of AIPs could be constructed simply on foundations of informative police cautions that truthfully communicate the implications of silence to suspects, rather than emphasise potential harm to their position, will be outlined. It will be contended that as both Scotland's disclosure regime and position on the right to legal assistance²² have now moved towards that of the rest of the UK,²³ the Review's stated aspiration that it seeks to avoid '(...) moving the trial process out of the courtroom and into the police station'²⁴ is ultimately untenable.

2. The Current Inconsistent Scottish Position

A. Adverse Inferences from Silence during Police Questioning

Walker and Walker state that: '(...) a person detained (...) is under no obligation to answer any question other than to provide certain specified information (...) so failure to answer other questions cannot be regarded as of evidential value.'²⁵ Ferguson reinforces this position, positing that '(...) there are no detrimental consequences in refusing to answer questions at a police station, since a trial judge

²⁰ The Carloway Review, *Responses to the Consultation Document* (n 4) 25.

²¹ *Murray v United Kingdom* (1996) 22 EHRR 29.

²² Criminal Procedure (Legal Assistance, Detention and Appeals) (Scotland) Act 2010, inserting an unqualified and universal right to legal assistance into ss 15 & 17 of the Criminal Procedure (Scotland) Act 1995. On the rationales justifying a right to legal assistance (which were also covered in her Consultation Response) see F Leverick, 'The right to legal assistance during detention' (2011) 15 (3) *Edinburgh Law Review* 352. Whilst the issue of waiver is also acknowledged here, word constraints prevent a full and frank discussion of the subject. See P Pleasance, V Kemp & NJ Balmer, 'The Justice lottery? Police station advice 25 years on from PACE' (2011) *Criminal Law Review* 3.

²³ Criminal Justice and Licensing (Scotland) Act 2010. Despite a clear nexus between administering cautions, AIPs and reciprocal disclosure obligations, there is insufficient space here to examine disclosure. For an excellent history, analysis and critique of Scottish disclosure requirements see F Raitt, 'Disclosure of records and privacy rights in rape cases' (2011) 15 (1) *Edinburgh Law Review* 33.

²⁴ The Carloway Review, *Report and Recommendations* (n 1) [7.5.24].

²⁵ Walker and Walker, *The Law of Evidence in Scotland* (3rd edn, Bloomsbury 2009) 104, referring to s14 (9) Criminal Procedure (Scotland) Act 1995.

must not invite the jury to draw adverse inferences from the suspect's silence.'²⁶ In terms of precedent offering support to these statements, the Review cites *Hoekstra v HM Advocate (No 5)*²⁷ wherein Lord Justice General Cullen found that a direction given to the jury in the lower court, in which it was indicated that the jury was permitted to draw an adverse conclusion about credibility from the accused's refusal to answer police questions, amounted to a misdirection.

It would therefore appear reasonable, on the basis of the authority just quoted, to assume that Scottish criminal procedure excludes the drawing of adverse inferences against an accused who has offered no responses during interview. Further critical analysis of the Review Team's published research somewhat muddies these waters, however. The Review initially states that: '[i]n relation to police questioning prior to trial, no adverse inference at all can be drawn from a failure to respond'.²⁸ They ground this prohibition on '(...) the antecedent caution, which expressly warns the suspect of [their] right not to answer questions.'²⁹ However the Review Team go on to observe that:³⁰

It is different if a person states something positive in response to an allegation and his/her answer, though not directly incriminating, implies some degree of involvement. What is not said in a response might be taken as meaning that the suspect accepts the allegation, or part of it, even although he/she does not say so expressly.

*Kay v Allan*³¹ is cited, with a footnote explaining that '(...) a partially answered question'³² would allow an inference to be drawn from what is not said in response to an allegation. This is confusing. It is unclear whether such an allegation must be made by police during questioning or by someone else at a different point in the criminal process. It also begs the question as to why an adverse inference could be permitted to be drawn from the silent part of a response (whatever that entails) but not from complete silence or 'no comment'. The Review then mentions 'ex tempore authority'³³ described as 'highly dubious'³⁴ that '(...) a statement by a co-accused in the presence of the accused is admissible evidence, but only in order to show the reaction of the accused to it.'³⁵ The case cited is *Buchan v HM Advocate*,³⁶ wherein

²⁶ PR Ferguson, 'Repercussions of the Cadder case: the ECHR's fair trial provisions and Scottish criminal procedure' (2011) 10 Criminal Law Review 743, 747.

²⁷ *Hoekstra v HM Advocate* 2002 SLT 599.

²⁸ The Carloway Review, *Report and Recommendations* (n 1) [7.5.5], citing *Robertson v Maxwell* 1951 JC 11.

²⁹ *ibid*. Despite the best efforts of this author, there appears to be no formalised wording for the Scottish police caution. The version allegedly taught at Tulliallan states: 'You are not obliged to say anything but anything you do say will be noted down and may be used in evidence' <<http://www.policespicals.com/forum/index.php?/topic/97298-does-anyone-know-the-scottish-police-caution/>> accessed 7 December 2012.

³⁰ The Carloway Review, *Report and Recommendations* (n 1) [7.5.5].

³¹ *Kay v Allan* 1978 SCCR Supp 188.

³² The Carloway Review, *Report and Recommendations* (n 1) [7.5.5].

³³ *ibid* [7.5.6].

³⁴ *ibid*.

³⁵ *ibid*.

³⁶ *Buchan v HM Advocate* 1993 SCCR 1076; 1995 SLT 1057.

Lord Justice-Clerk Ross made the following comments:³⁷

[T]he law regarding statements by persons other than the accused is to be found in *Lewis v Blair*³⁸ and is stated correctly in Renton and Brown's *Criminal Procedure* (5th ed), para 18-41a... "A statement by another person...made in [the] presence of an accused, is not in itself evidence against that accused. The accused's reaction to that statement, or indeed his failure to react to it where it is an accusation of his guilt, is, however, evidence against him in the same way as a statement made by him, silence in the face of accusation being capable of being construed as an admission of guilt. The evidence of the other person's statement is therefore admissible for the limited purpose of explaining the accused's reaction."

Two points are noteworthy here. Firstly, it appears that there are indeed occasions where Scots law will permit the drawing of adverse inferences from silence, albeit in somewhat confusing circumstances (such as the silent part of a partial response to a question,³⁹ or silence following accusation by a co-accused when made in the accused's presence).⁴⁰ Secondly, the fact that an express statement of the law, cited with approval in *Buchan*, is regarded by the Review Team as 'highly dubious' suggests that understanding and interpretation of this common law adverse inference 'provision' lacks consistency and clarity. Why silence in the aforementioned circumstances can be interpreted against an accused, whereas silence in response to police questioning (particularly since the advent of the right to legal advice prior to or during such questioning imposed after *Cadder*)⁴¹ cannot be interpreted in a similar fashion is inadequately explained. Such confusion lends weight to the argument for clear, consistent statutory AIPs.

B. Adverse Inferences from Silence at Trial

The Review's opening paragraph on silence at trial does little to clarify matters.⁴² It states that when an accused does not give evidence at trial, '(...) his/her silence cannot be used to prove, or assist to prove, fact, and, in that regard, to provide corroboration.'⁴³ Readers are then directed via footnote to compare this statement with *Maguire v HM Advocate*,⁴⁴ a case in which an accused's silence was in fact held to be corroborative. The preceding is not to be understood as an attack on the intellectual capabilities of the Review team, but rather as a demonstration that the current law is not being interpreted consistently. A further example of the inconsistency of the current Scottish position on the drawing of adverse inferences at

³⁷ *ibid* [1059]. The latest Renton & Brown *Criminal Procedure* (6th edn, 2013) still states this as the law at [24]-[56].

³⁸ *Lewis v Blair* (1858) 3 Irv 16.

³⁹ *Kay v Allan* (n 31).

⁴⁰ *Buchan v HM Advocate* (n 36).

⁴¹ *Cadder v HM Advocate* (n 2).

⁴² The Carloway Review, *Report and Recommendations* (n 1) [7.5.8].

⁴³ *ibid*.

⁴⁴ *Maguire v HM Advocate* 2003 SCCR 758; 2003 SLT 1307.

trial is provided by Walker and Walker, who state:⁴⁵

When the facts proved by the Crown raise an inference of guilt of the accused person, the accused bears a provisional burden of introducing contradictory facts or explanation. The burden arises purely through the state of the evidence and if the accused remains silent the court may well draw the inference least favourable to him.

The above suggests that the accused's silence becomes relevant if the proven facts before the court appear to indicate guilt. It is submitted that, in the interests of justice, facts should be placed before the accused much further back along the procedural chain (i.e. the police interview now taking place after legal advice) and that it should be possible to draw the 'least favourable' inference at this stage.

The Review cites a number of cases, including *Maguire, ante*, with a firm line of precedent dating back to *Hardy v HM Advocate*,⁴⁶ in which Lord Justice-Clerk Aitcheson opined that '(...) there are certain cases in which, in the absence of an explanation from the accused person, a jury may be amply entitled to draw an inference of guilt.'⁴⁷ It appears, therefore, that a clear and explicit common law AIP exists at the trial stage in Scotland. This common law AIP was applied in the terrorist appeal case of *McIntosh v HM Advocate*,⁴⁸ wherein Lord Justice-Clerk Ross opined that '(...) in relation to any inferences to be drawn (...) the trial judge was fully justified in reminding the jury that the appellant had not given evidence, and that (...) they might find it easier to draw from that evidence the inferences which the Crown invited'.⁴⁹ Citing *Donaghy v Normand*,⁵⁰ the Review goes on to state that '(...) the failure of an accused to testify, in circumstances where the evidence 'cries out' for an explanation, is a relevant factor which can be taken into account by a judge or jury when reaching a verdict'.⁵¹ It concludes by stating that '(...) a judge may comment on the failure of the accused to give evidence where the facts established by the evidence, if accepted, raise a *prima facie* inference of guilt'.⁵² The supporting cases cited are *Brown v Macpherson*⁵³ and *HM Advocate v Hardy*.⁵⁴ Yet having demonstrated that comment on silence at trial is acceptable, the Review's research then highlights the core flaw (namely inconsistency of interpretation) in not having a clearly defined statutory AIP framework, by observing that '(...) it has been made very clear by the courts that any such comment should be made with restraint and only in exceptional circumstances',⁵⁵ and further that '[f]or this reason, it is seldom done'.⁵⁶ The topic is

⁴⁵ Walker and Walker, *The Law of Evidence in Scotland* (n 25) [2.12.3].

⁴⁶ *Hardy v HM Advocate* 1938 JC 144.

⁴⁷ *ibid* [146].

⁴⁸ *McIntosh v HM Advocate* (No2) 1997 SLT 1320.

⁴⁹ *ibid* [1323]-[1324].

⁵⁰ *Donaghy v Normand* 1991 SCCR 877.

⁵¹ The Carloway Review, *Report and Recommendations* (n 1) [7.5.8].

⁵² *ibid*.

⁵³ *Brown v Macpherson* 1918 JC 3.

⁵⁴ *HM Advocate v Hardy* (n 46).

⁵⁵ The Carloway Review, *Report and Recommendations* (n 1) [7.5.8].

⁵⁶ *ibid*.

closed with the observation that since the Criminal Justice (Scotland) Act 1995⁵⁷ removed the prohibition whereby prosecutors could not comment adversely on an accused's failure to give evidence, it remains the fact that '(...) in practice, such comment is widely regarded as contrary to the spirit of the fair trial requirement and is very rarely made.'⁵⁸

It therefore appears that AIPs do exist in Scotland, however at a later stage in the criminal process than in other jurisdictions and in inconsistent circumstances. Nevertheless, their existence at the trial stage perhaps explains the Review's reluctance to introduce a system of codified AIPs, either in the belief that the existing provisions preclude the need or because to do so would be to move '(...) part of the trial out of the courtroom and into the police station.'⁵⁹ The circumstances under which adverse inferences are permitted at trial appear to be contingent on judicial precedent which significantly pre-dates the ECHR and which has been shown to be contradictory and confused. Prosecutorial comment also appears to be governed by an unwritten doctrine of 'fair play' that voluntarily prevents adverse inferences, even though the express prohibition was removed. The fact that the provisions already exist in their current common-law form makes the level of resistance to their introduction in statutory form difficult to understand. It appears that a culture of under-usage and ignorance of the true utility of the present position has developed, which is driving resistance to what in actuality would not be a significant change. It would not take a huge leap of faith to enshrine the current *ad hoc* provisions in an appropriate statute thereby rectifying this culture of resistance. It would also enable consistency in interpretation, particularly if the Scottish caution is adjusted to inform a suspect that adverse inference will be a possible outcome from any maintenance of contrived silence or refusal to comment. As will be argued later in this article, replacing the common law provisions with clear ECHR-centred cautions and a statutory framework governing the situations in which courts could draw adverse inferences not only allows for consistent (and therefore potentially more just) decision-making, but also sets out clearly to detained persons their responsibilities under law to account for themselves.

3. AIPs in Neighbouring Jurisdictions

As a member of the United Kingdom, Scotland is unique in not having AIPs codified in statute. In addition to making a case that argues that the Review was wrong to dismiss the idea of introducing such provisions in Scotland, it is appropriate to outline the AIPs as enacted in Northern Ireland, England and indeed in the Republic of Ireland.

A. Northern Ireland

The Criminal Evidence (NI) Order 1988 permits AIPs in certain circumstances.

⁵⁷ Criminal Justice (Scotland) Act 1995, s 32.

⁵⁸ The Carloway Review, *Report and Recommendations* (n 1) [7.5.9].

⁵⁹ *ibid* [7.5.24].

Article 3 sets out the circumstances whereby inferences may be drawn as a result of failure to mention particular facts when questioned by law enforcement agencies.⁶⁰ The caution associated with Article 3 expressly warns a person suspected of an offence, from the outset, that a failure to respond to relevant questions may have consequences.⁶¹ This caution is administered upon arrest and again prior to any interview. The Order provides for two further specific cautions for use as appropriate, each of which again expressly warns the detainee that adverse inferences may be drawn from their failure to adequately respond. These relate to requests to account for objects, substances or marks attributable to the commission of an offence (Article 5) and/or to account for their presence at a particular place (Article 6).

It is noteworthy that the 1988 Order has stood the test of time despite existing in a famously litigious environment. It has passed key tests of legitimacy and proportionality in the ECtHR.⁶² Such proven utility and Convention compliance provides a strong case for similar provisions in Scotland, albeit (as this article will suggest later) based on European Convention-centric, rather than ‘warning’ based, police cautions.

B. England and Wales

AIPs in England and Wales are set out in sections 34 to 39 of the Criminal Justice and Public Order Act 1994. Jackson observes that ‘[m]ost of the differences between the provisions of the 1994 Act and the original provisions in (...) Northern Ireland (...) no longer exist because the 1994 Act has amended the [NI] Order to conform almost entirely to the 1994 provisions.’⁶³ The English cautions replicate those in Northern Ireland. However, in contrast to their Northern Irish equivalent, the English provisions have been extensively criticised. This may be due to the fact that, in Northern Ireland, it is usually the judge personally (with no requirement to direct a jury) weighing the available proven evidence and deciding as to whether an adverse inference is appropriate. In England, on the other hand, it will be a jury that make the decision as to whether to draw the inference.⁶⁴ The Carloway Review explicitly referred to criticism of the English provisions in their ‘complexity’ argument, which is discussed later in this article. The English cautions are (as in Northern Ireland) in

⁶⁰ Criminal Evidence (NI) Order 1988, Article 3 (1).

⁶¹ Northern Ireland Office, *Police and Criminal Evidence (NI) Order 1989 Codes of Practice*, C10.5 <http://www.nio.gov.uk/police_and_criminal_evidence__ni_order_1989_codes_of_practice_2007_edition.pdf> accessed 16 March 2013.

⁶² *Murray v United Kingdom* (n 21).

⁶³ JD Jackson, ‘Interpreting the silence provisions: the Northern Ireland cases’ [1995] *Criminal Law Review* 587, 588. The 1994 Act’s amendments to the Criminal Evidence (NI) Order 1988 are in Schedule 10 [61].

⁶⁴ See, for example D Birch, ‘Suffering in silence. A cost-benefit analysis of section 34 of the Criminal Justice and Public Order Act 1994’ [1999] *Criminal Law Review* 769 and JD Jackson, ‘Silence and proof: extending the boundaries of criminal proceedings in the United Kingdom’ (2001) 5 (3) *International Journal of Evidence and Proof* 145.

daily use, with their interpretation governed by a judicial Specimen Direction.⁶⁵ This direction is guided by the ECtHR parameters set in the cases of *Murray*,⁶⁶ *Condrón*,⁶⁷ *Beckles*⁶⁸ and *Adetoro*.⁶⁹

C. Republic of Ireland

The Republic of Ireland, having observed the utility of AIPs in Northern Ireland and England, followed those jurisdictions by replacing the Criminal Justice Act 1984 with the Criminal Justice Act 2007. This Act permits Irish courts to draw adverse inferences from an accused's silence at both police interview and at trial.⁷⁰ Interestingly, the Carloway Review offered no criticism of the Irish provisions and there has been little in the way of academic criticism of the Irish AIP system. As in England, Wales and Northern Ireland, silence alone is incapable of grounding a conviction.

D. The Erroneous Preconception & Lessons from Other Jurisdictions

It is of critical importance to the debate regarding AIPs to emphasise several important facts. First, it is crucial to understand that detainees must be informed, in plain English, the meaning of each individual caution. The legislation in Northern Ireland expressly provides that it is the duty of the 'constable' to do so.⁷¹ Secondly, and of crucial significance, it should be emphasised that the mere act of remaining silent cannot, in any circumstances, ground a finding of guilt. There must be accompanying proved facts intimating guilt. This key safeguard is enshrined in the relevant judicial Specimen Direction.⁷² Finally, it is important to underline that an unfettered judicial discretion exists with respect to the use of AIPs. Otherwise put, in jurisdictions in which AIPs exist a court *may*, rather than *shall*, draw inferences when determining whether or not there is a case to answer or guilt at trial. A significant number of Consultation responses appeared to reflect the belief that adverse inferences, if introduced, would be drawn *as a rule* if an accused was brought to trial.⁷³ The law with respect to AIPs in Northern Ireland makes it clear that this is not the case. It is suggested here that this fundamental misconception about the discretionary nature of AIPs has, perhaps unwittingly, infiltrated the arguments of opponents to the introduction of AIPs in Scotland. This misconception has diverted

⁶⁵ Judicial Studies Board, *Crown Court Bench Book, Directing the Jury* (March 2010) (Specimen Direction (2004) 40.1) <http://www.judiciary.gov.uk/NR/rdonlyres/BE25EBB6-AAD2-4ACD-8115-28D3BF613164/0/benchbook_criminal_2010.pdf> accessed 19 March 2013.

⁶⁶ *Murray v United Kingdom* (n 21).

⁶⁷ *Condrón v United Kingdom* (2001) 31 EHRR 1.

⁶⁸ *Beckles v United Kingdom* (2003) 6 EHRR 162.

⁶⁹ *Adetoro v United Kingdom* [2010] ECHR 46834/06.

⁷⁰ See ss 18, 19 and 19a of the Criminal Justice Act 1984 as amended.

⁷¹ Criminal Evidence (NI) Order 1988, Article 5 (4).

⁷² See Judicial Studies Board for Northern Ireland, *Crown Court Bench Book* (3rd edition, 2010) 111-122 <<http://www.jsbni.com/Publications/BenchBook/Documents/BenchBook.pdf>> accessed 13 March 2013.

⁷³ Carloway Review, *Responses to the Consultation Document* (n 4). This tone permeates the responses of Ferguson, 25 and Leverick & Farmer, 34.

attention from the actual position outlined previously: that under statutory AIPs courts are actually empowered with an unfettered statutory discretion, enabling them to draw common-sense inferences only as a supplemental consideration, when evaluating other proven incriminatory evidence. The erroneous misconception that a statutory regime for AIPs would mean that inferences would always be drawn appears to have influenced the AIPs debate to a point where opponents may only be opposed because of their adherence to it.

If the above suggestion is accepted, it is further submitted that the question posed in the Consultation Document⁷⁴ represented a dice loaded (perhaps inadvertently) in favour of those who have adopted this misconception. It asked, '[s]hould the court be allowed to draw an adverse inference from a suspect's silence when questioned by police?'⁷⁵ This question immediately facilitates a presumption that blanket permission is being sought to draw adverse inferences in all cases in which an accused has maintained silence at interview, with no consideration afforded to the accompanying safeguards built into the legislation outlined above (such as appropriate cautions and the need for additional incriminating evidence). It is therefore inevitable that potential respondents would have been influenced towards adopting a more cautious and entrenched approach.⁷⁶ Even if this cannot be accepted, it is submitted that this question might have been more fairly phrased as follows:

In light of the right to legal advice now enshrined in Scottish law, and the enhanced disclosure requirements incumbent on prosecutors, should Scotland permit its courts, in circumstances in which other proven facts tend to demonstrate a suspect's guilt, to draw such inferences as appear sensible and proper from that suspect's silence during initial police questioning?

This question would have properly addressed the core issue and perhaps resulted in more measured responses. More weight may also then have been given to the responses supporting AIPs. If the AIPs debate is to be definitively decided on the basis of fact-based reasoning, this erroneous preconception must first be dispensed with and replaced with a clear understanding of the unfettered discretionary nature of statutory AIPs.

⁷⁴ Carloway Review, *Consultation Document* (n 3) 82.

⁷⁵ *ibid*, question 27. Question 28 invited respondents to speculate as to the impact of permitting such an inference. This did little to mitigate the effect of the preceding closed question and the conjecture invited would, without empirical evidence, do little to further this debate.

⁷⁶ See, for example the Carloway Review, *Responses to the Consultation Document* (n 4) 335, where the Scottish Society of Solicitor Advocates state: '[t]o allow such an inference would result in speculation and inevitably there would be a suggestion that the state is seeking to force a suspect to answer questions or give evidence thus diluting his right to a fair trial. It is unnecessary and unfair (...) [t]he likely scenario being: '(...) look at the evidence against him, he has not given evidence so he must be guilty.'

4. Roberts & Zuckerman and The Carloway Review

A. The Roberts and Zuckerman Factor

Referring to rationales sourced from Roberts and Zuckerman,⁷⁷ Leverick and Farmer's consultation response cites '(...) three principled reasons for recognising an absolute right to silence.'⁷⁸ The latter authors unknowingly played a significant role in underpinning the arguments of opposing respondents, with their analysis often selectively quoted to reinforce flawed reasoning. As an example of selective quotation, the aforementioned response of Leverick and Farmer appears somewhat misleading in referring to '(...) three principled reasons for recognising an absolute right to silence.' What Roberts and Zuckerman actually categorise are '(...) three types of argument (...) advanced as justificatory rationales for the privilege [against self-incrimination].'⁷⁹

Leverick and Farmer sidestep the fact that Roberts and Zuckerman caution that 'both the 'right of silence' and the 'privilege against self-incrimination' can be '(...) defined and disaggregated in different ways',⁸⁰ before finding '(...) fundamental inadequacies in the justificatory rationales supposedly underpinning the privilege.'⁸¹ Roberts and Zuckerman's ultimate conclusion is that '(...) none [of the rationales] individually or in combination, strikes us as particularly compelling.'⁸² At no point do they express support for an absolute right to silence, indeed they specifically caution that the '(...) juridical concepts [of a right to silence and the privilege against self incrimination] need to be disentangled and kept distinct.'⁸³

B. The Carloway Review's First Reason for Rejection

'The introduction of AIPs does not fit well with the presumption of innocence, the right to silence and the privilege against self-incrimination as understood and applied in Scotland.'⁸⁴ This reasoning amalgamates elements of all three possible justificatory rationales posited by Roberts and Zuckerman.⁸⁵ The flaws in this reasoning will shortly be demonstrated. Firstly, however, it should be underlined that the presumption of innocence as 'understood and applied in Scotland' is, for present purposes, understood as the idea '(...) that pre-trial procedures should be

⁷⁷ Roberts & Zuckerman, *Criminal Evidence* (n 14) 548. The authors identify three justificatory rationales for the privilege against self-incrimination: intrinsic (e.g. the protection of privacy and the prevention of cruel choices); conceptualist (such as adversarial procedure and the presumption of innocence); and instrumental (essentially the prevention against wrongful conviction).

⁷⁸ The Carloway Review, *Responses to the Consultation Document* (n 4) 40.

⁷⁹ Roberts and Zuckerman, *Criminal Evidence* (n 14) 548.

⁸⁰ *ibid* 540.

⁸¹ Roberts and Zuckerman, *Criminal Evidence* (n 14) 579.

⁸² *ibid*.

⁸³ *ibid* 539. They specifically state that 'the privilege against self-incrimination' is often invoked loosely and has a tendency to become confused with the overlapping notion of the 'right to silence.' In their response to the consultation, it appears that Leverick & Farmer have fallen victim to this tendency. See Carloway Review, *Responses to the Consultation Document* (n 4) 34.

⁸⁴ The Carloway Review, *Report and Recommendations* (n 1) [7.5.26].

⁸⁵ Roberts and Zuckerman, *Criminal Evidence* (n 14) 548.

conducted, so far as possible, *as if* the defendant were innocent.⁸⁶ Similarly, for present purposes, 'right to silence' and 'privilege against self-incrimination' comprise the right to silence as outlined by Lord Mustill,⁸⁷ juxtaposed with the ECtHR's seminal findings in *Murray*;⁸⁸ namely that the '(...) right to remain silent under police questioning and the privilege against self-incrimination are generally recognised international standards which lie at the heart of Article 6.'⁸⁹

Of more importance in advancing this article's case is the following statement contained in the ECtHR's *Murray* judgment, which it is necessary to restate in full:⁹⁰

What is at stake (...) is whether these immunities are absolute, in the sense that the exercise by an accused of the right to silence cannot under any circumstances be used against him at trial or, alternatively, whether informing him in advance that, under certain conditions, his silence may be used, is always to be regarded as "improper compulsion." On the one hand, it is self-evident that it is incompatible with the immunities under consideration to base a conviction solely or mainly on the accused's silence or on a refusal to answer questions or to give evidence himself. On the other hand, the Court deems it equally obvious that these immunities cannot and should not prevent that the accused's silence, in situations which clearly call for an explanation from him, be taken into account in assessing the persuasiveness of the evidence adduced by the prosecution.

Wherever the line between these two extremes is to be drawn, it follows from this understanding of "the right to silence" that the question whether the right is absolute must be answered in the negative. It cannot be said therefore that an accused's decision to remain silent throughout criminal proceedings should necessarily have no implications when the trial court seeks to evaluate the evidence against him.

Although it is appreciated that European and English jurisprudence regarding AIPs has evolved significantly since *Murray*, these original findings have never been overruled. Strasbourg fully endorsed the Criminal Evidence (NI) Order 1988 provisions. More specifically, Strasbourg endorsed the findings of the European Commission on Human Rights, agreeing that:⁹¹

(...) the provisions (...) constitute a formalised system which aims at allowing common sense implications to play an open role in the assessment of evidence. The Commission finds no indication on the facts of this case that it

⁸⁶ See A Ashworth, 'Four threats to the presumption of innocence' (2006) 10 (4) International Journal of Evidence and Proof 2006, 243.

⁸⁷ *R v Director of Serious Fraud Office, ex parte Smith* [1993] AC 1 [30-31]. Here, Lord Mustill defined the right to silence as '[not denoting](...) any single right, but rather...a disparate group of immunities, which differ in nature, origin, incidence and importance...' Please see case for Lord Mustill's full elaboration of what 'the right to silence' entails.

⁸⁸ *Murray v United Kingdom* (n 21).

⁸⁹ *Murray v United Kingdom* (n 21) [45].

⁹⁰ *ibid* [47].

⁹¹ *ibid* [54].

deprived the applicant of the right to silence or that the consequences which flowed from his exercise of that right were unfair.

This statement by Europe's highest court immediately renders the Review's 'main argument'⁹² untenable. As AIPs are plainly afforded legitimacy by Europe's highest court (who will have weighed these conceptual arguments more often and in greater depth than the Review team), it is disappointing that the Review team used such demonstrably unconvincing conceptual analysis as its principal reason for rejection. If Europe's highest court has sanctioned the use of AIPs, any legal argument made opposing their introduction in any European jurisdiction begins to lack cogency. Furthermore, it is difficult to understand the logic of opponents when the analogous situation surrounding the right to legal assistance is considered. In *Cadder*, the Supreme Court's application of the European Court's findings in *Salduz*⁹³ compelled the introduction of legislation finally bringing Scotland into line with a European standard her courts had fought tooth and nail to evade. Examination of the associated case law appears to suggest that Scottish judicial resistance was founded mainly on a perceived need to protect cherished historic procedural differences, rather than any aspiration towards ensuring a Convention-compliant legal system.⁹⁴ The ultimate futility of such a stance was rather embarrassingly brought home in *Cadder*. If Europe's highest court requires a right to legal assistance prior to questioning, and Scotland's government feels forced to comply, where is the credibility in refusing to pre-emptively update Scottish investigative procedure with cautions and provisions legitimised by the same court?

A second significant point of concern is that this principal reason for rejection appears to have been arrived at by the Review team in isolation, with little in the way of support from the responses that they were supposed to consider. Detailed scrutiny of the responses uncovered only two which expressly alluded to this reasoning, namely those of the Scottish Liberal Democrats⁹⁵ and Justice.⁹⁶ Cynics could be forgiven for concluding that greater weight was given to the views of a political (as opposed to a legal) entity (which of itself could be construed as partisan) and those of a civil rights organisation, rather than to the views of Scotland's principal law reform body,⁹⁷ its rank and file police service⁹⁸ and victim support groups.⁹⁹ It appears that as well as ignoring leading European jurisprudence on the subject, the Review ignored virtually all of its own consultation responses in arriving at this reason for rejection. This does not inspire confidence that the subject has been approached in a balanced and comprehensive manner.

⁹² The Carloway Review, *Report and Recommendations* (n 1) [7.5.23].

⁹³ *Salduz v Turkey* (2009) 49 EHRR 19.

⁹⁴ For an excellent synopsis of the background to Scotland's forced conversion to a right to legal assistance see F Leverick 'The right to legal assistance during detention' (n 22).

⁹⁵ The Carloway Review, *Responses to the Consultation Document* (n 4) 156.

⁹⁶ *ibid* 217.

⁹⁷ *ibid* 200.

⁹⁸ *ibid*. See the Scottish Police Federation at 87, the Association of Scottish Police Superintendents at 102 and the Association of Chief Police Officers in Scotland at 250. All police responses supported AIPs.

⁹⁹ *ibid*. Particularly those representing the rights of rape victims. See Rape Crisis Scotland at 163.

Returning to academic grounds to finally and convincingly rebut this particular aspect of the Review team's reasoning, it was suggested earlier that this reasoning amalgamated elements of each of Roberts and Zuckerman's justificatory rationales, in particular the 'conceptualist rationale.'¹⁰⁰ The latter authors define the 'conceptualist rationale' as one that includes the assertion that the privilege against self-incrimination '(...) follows by necessary implication from the presumption of innocence.'¹⁰¹ Correctly observing that '(...) conceptual analysis can never be an adequate surrogate for moral argument in rationalizing criminal procedure',¹⁰² Roberts and Zuckerman describe three versions of the argument markedly akin to that underpinning the Review's first reason for rejection. Alternatively phrased versions of this argument also appear in several of the Consultation Responses.¹⁰³ This section will now outline each version, followed by the authors' considered and logical rebuttals.

Firstly, there is the argument that '(...) removing the privilege against self-incrimination would diminish the prosecution's duty to prove guilt beyond reasonable doubt.'¹⁰⁴ The authors rebut this line of reasoning by stating that it:¹⁰⁵

(...) confuses a normative standard of sufficiency of proof with the type of evidence capable of satisfying that standard. Where silence is evidentially probative it may contribute towards discharging the prosecutor's burden of proof without in any way diluting the traditional standard of proof (...) there is no more justification for preventing the prosecutor from relying on probative silence, than for discounting other apparently incriminating evidence, such as for example (...) that the suspect was seen fleeing from the scene of the crime.

Secondly, there is the argument that '(...) the privilege is a component of the accused's right not to have to (...) defend himself unless the prosecution first establishes a *prima facie* case.'¹⁰⁶ Again, Roberts and Zuckerman effectively negate this assertion by observing that:¹⁰⁷

Even if the privilege were abolished the prosecution should still continue to bear the burden of establishing a *prima facie* case (...) silence in court might *contribute* to a *prima facie* case, but could never constitute a case to answer. Silence (...) is only ever probative in the context of other incriminating evidence.

This statement reinforces the factual position as set out in the Northern Irish and English legislation, namely that silence of itself does not automatically permit the

¹⁰⁰ Roberts and Zuckerman, *Criminal Evidence* (n 14) 540.

¹⁰¹ *ibid* 554.

¹⁰² *ibid*.

¹⁰³ The Carloway Review, *Responses to the Consultation Document* (n 4). See C Christie at 7, F Stark at 44, Scottish Liberal Democrats at 156, Justice at 217 and the Senators of the College of Justice at 359.

¹⁰⁴ Roberts & Zuckerman, *Criminal Evidence* (n 14) 554.

¹⁰⁵ *ibid*.

¹⁰⁶ *ibid*.

¹⁰⁷ *ibid* 554-555.

drawing of an adverse inference. Roberts and Zuckerman's third version of the argument underpinning that relied upon by the Review is as follows:¹⁰⁸

(...) it might be contended that the presumption of innocence is greatly weakened if accused persons are required to account for themselves during police interrogation. The prosecution might then rely on (...) suspicious silences in order to establish a *prima facie* case at trial. This (...) is all the more perplexing to the extent that police are not obliged to disclose information (...) prior to an interrogation, beyond the bare fact that he is suspected of a particular offences (...) the presumption of innocence is truly a bulwark of autonomy, a cipher of justice and a servant of democratic accountability in liberal societies.

The authors rebut this effectively by correctly observing that powers of arrest, detention and right to question persons in custody are set out in statute '(...) completely independently of the privilege (...) [which] itself consequently does little to safeguard the citizen from the perils of arrest, detention and police interrogation.'¹⁰⁹ Put bluntly, the existence of a privilege against self-incrimination, already disentangled from an absolute right to silence, in no way protects a person from being accused of a crime, arrested and brought to a police station and interviewed with a view to eliciting their account of the circumstances which brought about their arrest or detention. However, the privilege as understood in Scotland currently hands the suspect an early advantage in obstructing the pursuit of truth, whereby the original grounds for arrest are neutered by a caution which tells the suspect that they are under no obligation to assist the investigative process at all. This state of affairs has been compounded since the advent of a right to legal assistance in Scotland as the accused can legally contrive to remain silent. As will be seen shortly, the above approach is at odds with less controversial powers occasionally employed to obtain non-verbal evidence forcefully, such as samples, which of themselves can lead to adverse inferences being drawn.

Some of the most vehement opposition to AIPs was expressed in the consultation response of Ferguson, who has previously described AIPs as a '(...) potentially legitimate form of compulsion.'¹¹⁰ She asserts that:¹¹¹

(...) when failure to respond to police questioning is treated as an indication of guilt, there is no true "right" to silence at all: silence is being treated as a form of self-incrimination. We tend to think of compulsion as coming from physical or psychological threats, but it may equally come from the knowledge that one's silence may form part of the prosecution's case.

It is respectfully submitted that Ferguson's view is too strong and fundamentally inaccurate. Failure to respond to police questioning is not treated as an indication of guilt *any more* than the original arrest upon reasonable suspicion or complaint is

¹⁰⁸ *ibid* 555.

¹⁰⁹ Roberts & Zuckerman, *Criminal Evidence* (n 14) 555.

¹¹⁰ Ferguson (n 26) 753.

¹¹¹ Ferguson (n 26) 753-754. This exact statement was reproduced in her response (Carloway Review, *Responses to the Consultation Document* (n 4) 29.

treated as an indication of guilt. In jurisdictions with AIPs, failure to respond to police questioning (after a suitable caution and the benefit of legal advice delivered prior to interview) *may*, in conjunction with (*and only in conjunction with*) other proven evidence, be subject to an adverse inference. Silence *is not* treated as a form of self-incrimination; rather silence *may* simply be viewed as an additional contributory indicator of participation in the offence(s) alleged when assessed alongside other incriminating evidence. It is impossible to incriminate oneself through maintaining silence. Incrimination only arises through the inferences that can be drawn from all the other available evidence, with a failure to adequately explain the presence of that evidence being permitted to give additional weight to all the proven facts. It is submitted here that there is no compulsion element whatsoever. Ferguson's contentions conform to Roberts and Zuckerman's 'intrinsic rationales',¹¹² whereby remaining silent allegedly '(...) spares the accused the 'dilemma' of the choice between testifying truthfully and contributing to his own conviction, or committing perjury in order to escape punishment.'¹¹³ These authors argue that none of the '(...) modern rationales emphasizing the hardship of compelled self-incrimination (...) satisfactorily neutralises Bentham's original critique of the privilege against self-incrimination.'¹¹⁴ Ferguson's arguments regarding privacy and personal autonomy also mirror this intrinsic rationale. She contends that:¹¹⁵

The suspect (...) is in a vulnerable position. S/he may be unaware of the full nature of the allegations being made against him/her and the basis for these allegations. On (*sic*) should not be required to account for what one was doing/where one was/who one was with at such a preliminary stage in the proceedings.

However, Roberts & Zuckerman correctly observe (with reference to Article 8 of the ECHR) that:¹¹⁶

(...) the right to privacy is subject to numerous qualifications and limitations (...) although people should be left alone with their private thoughts (...) it is not oppressive, where material suspicions of offending are aroused, to invite citizens to respond to accusations or to account for apparently incriminating circumstances (...) the right to privacy is no more persuasive justification for silence in the police station than at court.

As regards the idea that a detainee should not have to account for their movements, or for objects or marks, it is argued that, in the interests of pursuing truth and justice, this is exactly what they should have to do. Investigative agencies are required to obtain evidence and process accused persons as swiftly and diligently as possible. Ferguson argues that there should be adequate time for a suspect to consider advice and to form a response to the allegations made against him/her. It is

¹¹² Roberts & Zuckerman, *Criminal Evidence* (n 14) 549.

¹¹³ *ibid.*

¹¹⁴ *ibid* 550.

¹¹⁵ The Carloway Review, *Responses to the Consultation Document* (n 4) 29.

¹¹⁶ Roberts & Zuckerman, *Criminal Evidence* (n 14) 551-552.

submitted here that the time between initial suspension of liberty and the receipt of the prescribed legal advice is more than adequate to form at least the basis of a defence. Some of the evidence the police are required to secure expeditiously is best obtained by questioning, which now cannot take place until after legal advice has been received. Any decision to remain silent requires a conscious effort to resist questions designed to elicit guilt or innocence and impacts directly on the quality of evidence being retrieved. In turn, this impacts on the quality of the evidence that can be placed before a court, meaning that the state's duty to protect society by bringing suspected offenders before that court is obstructed. Whether or not an accused is aware of the full nature of the allegations being made, it cannot seriously be countenanced that total silence is not a contrived condition.

Ferguson expresses support in her response for the idea of a distinct legal status of 'suspect' with associated rights and permitted infringements.¹¹⁷ She has no issue with a court drawing adverse inferences from a suspect's failure to voluntarily submit to an identification parade or from a suspect's failure to co-operate with the taking of samples.¹¹⁸ This is despite the fact that obtaining such samples may require police or medical staff to legally assault a non-compliant suspect, something which is unthinkable when contemplating how to elicit verbal evidence. If adverse inferences appear permissible despite these infringements of the accused's body, it is difficult to understand the objection to drawing an inference from contrived silence, where the accused has received legal advice, his/her interview is audio/video recorded and where the contents of his or her mind are of at least as much evidential value as other physical evidence.

In sum, it is therefore submitted that the Review's first published reason for rejecting AIPs fails to accord sufficient weight to European Court jurisprudence on the subject (namely the paradigm case of *Murray*, which still stands), and inflates the importance of arguments based on selectively quoted Roberts and Zuckerman potential rationales, which in fact have all been analysed and discredited by these authors themselves. Such reasoning simply does not stand up to scrutiny.

5. The Carloway Review's Second Reason for Rejection

As noted in the introduction to this article, the second argument advanced by opponents of AIPs in the Review is that AIPs would bring unnecessary complexity to the criminal justice system. This argument for rejecting AIPs is potentially stronger than the principled argument outlined in the part above and directly opposes this article's stated position that statutory AIPs would bring clarity and consistency to Scottish criminal procedure. It is a practical, seemingly logical argument which does not fit the aforementioned rationales for rejection. It also featured in many of the Consultation responses.¹¹⁹ In this respect at least this reason for rejection is reflective

¹¹⁷ The Carloway Review, *Responses to the Consultation Document* (n 4) 25.

¹¹⁸ *ibid.*

¹¹⁹ *ibid.* 8 of the 19 responses opposing AIPs can be read as opposing AIPs on, *inter alia* 'complexity' grounds. See for example Henderson at 19, Leverick & Farmer at 34, Stark at 44 and Chalmers at 57.

of the views of respondents. It can however be rebutted, and in doing so, two complementary arguments are offered.

Firstly, it is argued here that in citing the problems and complexity associated with English AIPs as a fundamental reason not to adopt such provisions in Scotland,¹²⁰ little or no consideration was afforded to the idea that Scotland does not have to follow the English model, but can instead draft its own AIPs. Suggested cautions are set out below further below. Redmayne (also selectively cited by opponents of AIPs) points out that '[e]xporting section 34 [i.e. its potential consequences] would not necessarily mean exporting the case law that has accreted to it.'¹²¹ It is conceded that English AIPs have become complex and convoluted, due principally to the machinations of successive judges rather than the original drafting. The English cautions, which refer to 'warning' and 'harming a defence', are inherently unsuitable when set against the aspiration of an overall Convention-compliant 'fair trial' process. Hamilton (in assessing the current state of the presumption of innocence in Ireland) has said the following about the English system:¹²²

(...) the much more structured approach taken by the English courts means that a comparison with recent English jurisprudence is unfavourable. This divergence [in how AIPs are applied and interpreted] is perhaps surprising given the recent incorporation of the European Convention (...) in both jurisdictions.

This statement acknowledges not only that Ireland's courts have not felt compelled to develop similar convoluted interpretations of their AIPs, but also, more crucially, that a small Member State utilising a similar adversarial trial system as Scotland, within an evolving Area of Freedom, Security and Justice,¹²³ can and does use the margin of appreciation available to it to operate Convention-compliant AIPs.¹²⁴ Scotland, with a blank canvas currently available, can do likewise.

As mentioned above, the English caution refers to 'harming a defence'. The Northern Ireland caution also uses this language. In pursuit of a just criminal procedure and a fair trial, it may perhaps be better to concentrate on advising a person in custody (who has received or is receiving legal advice) that, if they wish to express their right to silence or not offer any explanatory words in response to questioning, that a court may, when considering all the relevant proven evidence, also take a view of their failure to account for themselves. This author is of the view that the wording contained in the English and NI models, in the context of delivery

¹²⁰ *ibid.* Leverick & Farmer's response, at 34, explicitly states that they have '(...) drawn much assistance from AIPs in England & Wales.'

¹²¹ See M Redmayne, 'English Warnings' [2008] *Cardozo Law Review* 1047, 1086. Although critical of the English provisions, Redmayne does not adopt a position of being against AIPs *per se*.

¹²² C Hamilton, 'Threats to the presumption of innocence in Irish criminal law: an assessment' (2011) 15(3) *International Journal of Evidence and Proof* 181 at 205.

¹²³ Consolidated Treaty on the Functioning of the European Union, Title V, Area of Freedom, Security and Justice. See in particular Chapter 4, Judicial Cooperation in Criminal Matters and Chapter 5, Police Cooperation in Criminal Matters.

¹²⁴ Hamilton (n 122) cites no challenges to the Irish AIPs, either in Europe or at Irish Supreme Court level. Furthermore the Review's synopsis of the Irish provisions makes no criticisms of them.

to a vulnerable detainee (from any authority figure) provides ammunition to opponents of AIPs seeking to legitimately assert the existence of a degree of compulsion.

On another note, although 'complexity' is sometimes a cause for judicial 'anxiety',¹²⁵ criminal law is by definition, inherently complex. 'Complexity' is a subjective term. Whilst it featured prominently as a potential problem with AIPs in the responses of opponents, no responses in support expressed similar concerns. Complexity is in the eye of the beholder and consequently can be interpreted differently by supporters and opponents of AIPs. The Review's complexity argument may, in reality, reflect an unwillingness to embrace necessary change or a desire to maintain cherished 'differences'. It may alternatively be founded on a perception that educating Scotland's police, prosecutors, the legal profession and judiciary as to AIPs represents too burdensome or costly a task.¹²⁶ The complexity rejection can be easily satirised to read that cost and bureaucratic/logistical reservations do not fit well with the Article 6 right to a fair trial as understood in Scotland. What may be currently missing in Scotland is a sufficient impetus or collective will on the part of the legal establishment to adapt to the evolving European model (to which Scotland is inextricably, if indirectly bound).¹²⁷ The Review itself alluded to this requirement for the legal profession to adapt.¹²⁸

The preceding argument in no way implies that Scotland should homogenise its procedures with those of Europe, but merely that harmonization may be the way forward.¹²⁹ It is worth underlining, however, that '(...) continental/civilian jurisprudence has never found any difficulty in condoning common sense inferences from silence.'¹³⁰ Given that most of Europe has little difficulty with the concept of AIPs (albeit due to the inquisitorial trial model) there remains little justification for Scotland to continue to exclude them. It is submitted that the dismissal of AIPs on grounds of undue 'complexity' uses irrelevant English based evidence, ascribes an overly harsh subjective interpretation as to what constitutes 'complexity' and fails to grasp the reality of Scotland's current position (and future direction) within Europe.

6. Making it Happen

The key to introducing a Convention-compliant criminal procedure facilitating AIPs lies in the introduction of new statutory cautions.¹³¹ Having outlined the

¹²⁵ See the comments of the judge in *R v Bresa* [2005] EWCA Crim 1414, cited by Ferguson in her consultation response (Carloway Review, *Responses to the Consultation Document* (n 4) 29).

¹²⁶ See the analogous view of Roberts and Zuckerman, *Criminal Evidence*, (n 14) 560, observing, in relation to England, that '(...) the duty solicitor scheme (...) boasts impressive geographical coverage, but such expenditures are not obvious vote-winners.'

¹²⁷ An example of the evolving European model is the European Commission's Proposal for a Directive on the Right of Access to a Lawyer etc (COM(2011) 326 final) <<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0326:FIN:EN:PDF>> accessed 12 July 2012.

¹²⁸ The Carloway Review, *Report and Recommendations* (n 1) [5.2.31].

¹²⁹ On this debate see Ferguson (n 26) 756.

¹³⁰ Roberts and Zuckerman, *Criminal Evidence* (n 14) 579.

¹³¹ Carloway Review, *Responses to the Consultation* (n 4). See also the views of Raitt, at 61 and the Association of Scottish Police Superintendents at 102 who support this position.

shortcomings of the English caution(s) and dismissed the relevance of referring to the English system when contemplating a scheme for Scotland, suggested wording for an initial caution upon detention is suggested by way of footnote.¹³² Prior to questioning, and after legal advice has been received or waived (preferably with a legal representative present) a suspect could then be formally cautioned as per the footnote below.¹³³ Whilst this caution appears lengthy in comparison to its UK counterparts, potential criticism is countered by the fact that it could be in print and available in relevant custody areas, with the suspect invited to retain a signed copy. It represents a clear statement of the suspect's Convention rights and reciprocal responsibilities. It carries no intimation of compulsion. As the written caution constitutes a disclosable document, there will be no issue for the defence or the courts in proving that the investigating team have adhered to the need for Convention fairness. For those unable to read the document, the caution could be fully explained (on tape) in layperson's terms prior to interview, with an indication sought from the detainee as to their comprehension of it. This article submits that it is possible for a ECHR-compliant procedure, which commences in every other respect when a person is first detained, to permit that detainee to be clearly informed not only of their rights, but of an obligation to verbally provide their explanation that can then be assessed in conjunction with any other evidence. The cautions as suggested ensure that that obligation is fairly explained to them.

7. Conclusion

It has been shown in this article that AIPs do currently exist in Scotland, but on an inconsistent basis, combining occasionally conflicting judicial precedent and an unwritten doctrine of fair play practiced by prosecutors. The relevant AIPs in Northern Ireland, England and the Republic of Ireland have been briefly described, along with a reminder to opponents of the fact that the ECtHR fully legitimised such provisions in *Murray*. This fact renders the Review's first 'conceptual' reason for

¹³² Your Convention right to liberty is now suspended, as I suspect you of involvement in the offence(s) of _____. This means that you are now arrested/detained. You will be brought to _____ (Police Station or other place of detention). You have a right to silence, under the European Convention on Human Rights and anything you do say to me now will be recorded and included in my evidence. At _____ (place of detention) you have an immediate Convention right to a private consultation with a solicitor before I/we pursue our investigation any further.

¹³³ You are reminded as to why your Article 5 right to liberty has been suspended; namely suspicion of involvement in _____. You have received/are receiving professional legal advice in relation to that/those matters. You will now be questioned by us in relation to that matter as part of our investigations. You have a right to remain silent or not to answer any of these questions under the European Convention on Human Rights. However the European Court of Human Rights has held this right not to be absolute. All of your responses, or the fact that you remained silent will be (tape) recorded. If this matter proceeds to a hearing in a court, that record will be presented as part of our evidence. Because you also have a Convention right to a fair legal process, you are strongly advised to tell us during this interview, of any fact or explanation you may have to offer which may tend to assist in proving your innocence. If you do not take the opportunity during this interview to do so, a court may, when assessing all the evidence we put forward relating to you, infer that you, by not taking this opportunity to offer any fact or explanation, have no innocent explanation to offer.

rejecting Scottish AIPs untenable. The main opposing arguments (analysed by reference to their selectively quoted Roberts and Zuckerman potential rationales) have also been rebutted. It is revealing that no opposing respondent, or indeed the Review team, acknowledged Roberts and Zuckerman's ultimate conclusion that '(...) the interests of suspects in the police station, and of the accused at trial, are more fairly and effectively protected by tailor-made procedural safeguards, such as those set out in minute detail by PACE 1984 and its associated Codes of Practice.'¹³⁴ This article fully endorses that conclusion.

The Review's second (practical) argument regarding potential complexity (relying on the prevailing complexity of the English provisions) has been shown to be irrelevant in contemplating AIPs for Scotland, being more likely motivated by perceptions of cost and logistics. Law is, by definition, complex, and the term is subjective. There is little complexity in the cautions proposed herein, which do not represent finalised versions, but an embryonic way forward. Furthermore, if criminal law is continually 'tweaked', rather than repealed and renewed in line with prevailing societal needs, it will inevitably become increasingly complex. The Criminal Procedure (Scotland) Act 1995 is living testament to this.¹³⁵ It is suggested that the published Review as a whole should be referred to the SLC with a view to creating a single complete statute on Criminal Investigative Procedure for Scotland, similar, but not necessarily identical to the Police and Criminal Evidence Act 1984. Despite predating the Scottish Act by eleven years, the Act's provisions have stood the test of time more steadfastly. This would be preferable to the current piecemeal reforms coming forth in relation to legal access and disclosure which, although welcome, mean that various statutes have to be referred to rather than a single PACE equivalent law and associated Code of Practice. This in itself may create undue complexity.

There can be no doubt that the Carloway Review represents a commendable achievement when set against the magnitude of the task it was presented with. It is disappointing, however, that it used such demonstrably weak reasoning in its rejection of AIPs. This is compounded by the fact that the responses of those supporting AIPs for Scotland were totally ignored in the published document. All parts of Scotland's police service supported AIPs. As the public service at the coal face of investigating reported crime, their views should have carried greater weight. Despite the well-documented obstacles to securing justice for Scotland's rape victims, their views (and other victims' rights groups) also carried little weight. It is strongly suggested that with SLC interest in re-examining the matter,¹³⁶ Scottish Ministers should refer the topic to the Commission in order for the law relating to adverse inferences to receive the level of scrutiny it deserves. Victims of crime need to know that every Convention-compliant procedure is available to agencies in Scotland charged with investigating crime.

¹³⁴ Roberts and Zuckerman, *Criminal Evidence* (n 14) 579.

¹³⁵ There have been over 120 modifications to the original Act.

¹³⁶ Carloway Review, *Responses to the Consultation Document* (n 4) 200.

The Roma Minority & Free Election Rights in Bosnia and Herzegovina: The Problem of Constitutional Reform

JENNIFER PATON*

Abstract

The Constitution of Bosnia and Herzegovina ('BiH') contains provisions that discriminate against minority groups by depriving them of the right to participate in free elections without distinction on the basis of ethnicity. This is a right enshrined in international and regional human rights instruments by which the state is bound. In failing to afford minority groups the right to participate in free elections, BiH is failing to comply with its obligations under these instruments. By focusing on the plight of the Roma minority in BiH, it will be contended in this paper that urgent constitutional reform is required to end this constitutionally-mandated discrimination and the detrimental impact it has on the Roma minority. Furthermore, it will be demonstrated that the persistent failure of the authorities in BiH to affect such reform now threatens BiH's membership in the Council of Europe and its progress towards membership in the European Union. It will be concluded that BiH must overcome the difficulties inherent in achieving the necessary constitutional reform prior to the 2014 general elections, or risk political isolation from its European neighbours.

1. Introduction

The Constitution of Bosnia and Herzegovina¹ was drafted with the principal aim of restoring peace and stability in the state following the Bosnian War.² However, this aim was pursued at the expense of the rights of minority groups. This article will focus specifically on the Roma minority and the right to free elections. It will look, first, at the Constitution in the context of post-war BiH, as well as at the social and political position of the Roma in the state following the end of the conflict. It will then address the protection afforded to human rights in BiH, specifically the right to participate in free elections. It will be argued that the Constitution continues to contain provisions that discriminate against minorities by denying them to right to participate in free elections without distinction on the basis of their ethnicity. This article will then consider whether BiH's recent turbulent history can justify such discrimination. Ultimately, this article will demonstrate that there is a pressing need

* Graduate, School of Law, University of Aberdeen.

¹ Annex 4, *General Framework Agreement for Peace in Bosnia and Herzegovina* ('Dayton Agreement'), 21 November 1995 <<http://www.unhcr.org/refworld/docid/3de495c34.html>> accessed 9 March 2013 (henceforth referred to as 'the Constitution').

² Preamble to the Constitution, *ibid.*

for constitutional reform in order to bring BiH into line with its international and regional obligations regarding state policy towards the Roma minority. However, there are concerns about the extent to which constitutional reform can address the causes and detrimental effects of discrimination against the Roma, and this issue will also be discussed. This article will conclude that the potential consequences of continued inactivity by the authorities of BiH go beyond the detrimental impact of discrimination on the Roma minority: BiH's membership of the Council of Europe is at risk, and its progress towards membership of the European Union has stalled. Both the Council of Europe and the European Commission have indicated that further general elections held in accordance with the discriminatory constitutional provisions will be unacceptable to both organisations.³ In these circumstances, there is an urgent need for constitutional reform enabling the Roma minority to enjoy the right to free elections without discrimination.

2. Post-War BiH and the Roma Minority

A. The Constitution of BiH

It is necessary to briefly examine the background against which the Constitution of BiH was drafted and to set out the pertinent constitutional provisions. The Constitution is annexed to the Dayton Agreement, which was negotiated by the Republic of Bosnia and Herzegovina, the Republic of Croatia and the Federal Republic of Yugoslavia at Dayton, in the USA, under the supervision of international mediators. It was signed on the 15th December 1995, ending the Bosnian War which followed the break-up of the Socialist Federative Republic of Yugoslavia. The belligerent parties in this war were the Bosniacs, Serbs and Croats⁴ and it was clear before and during the time of drafting that the resulting Constitution would have to protect and balance the competing rights and interests of these three groups if enduring peace and stability were to be ensured. This was achieved by giving the three groups the status of 'constituent peoples'.⁵ Affiliation with a constituent people is by self-identification.⁶ The Preamble to the Constitution states that all those who do not identify with one of the constituent peoples are 'others'.⁷

The Constitution establishes the federal state of Bosnia and Herzegovina, comprised of two entities: the Federation of Bosnia and Herzegovina, where Croats and Bosniacs are predominant, and the Republika Srpska, where Serbs are predominant.⁸ The state exercises its powers by means of a series of institutions, the Parliamentary Assembly and the Presidency being the most significant for the purposes of this article. The Parliamentary Assembly is a bicameral assembly

³ European Commission, 'Statement by Commissioner Štefan Füle on political consultations on implementation of the Sejdić-Finci judgement' MEMO/13/328, 11 April 2013.

⁴ S Bardutzky, 'The Strasbourg Court on the Dayton Constitution: A judgement in the case of *Sejdić and Finci v Bosnia and Herzegovina*' (2010) 6(2) ECL Review 309, 310.

⁵ Preamble to the Constitution (n 1).

⁶ *Sejdić and Finci v Bosnia and Herzegovina* 28 B.H.R.C. 201 [11] (hereinafter '*Sejdić*').

⁷ Preamble to the Constitution (n 1).

⁸ Constitution (n 1) Article I (3).

comprised of a House of Representatives and a House of Peoples.⁹ Legislation must receive the approval of both chambers.¹⁰ The House of Peoples must be comprised of '(...) 15 Delegates, two-thirds from the Federation (including five Croats and five Bosniacs) and one-third from the Republika Srpska (five Serbs)'.¹¹ Although an individual must affiliate with one of the constituent peoples to become a delegate to the House of Peoples, there is no such restriction on election to the House of Representatives.¹² The three-person Presidency is established by Article V, which states:

The Presidency of Bosnia and Herzegovina shall consist of three Members: one Bosniac and one Croat, each directly elected from the territory of the Federation, and one Serb directly elected from the territory of the Republika Srpska.

Again, membership of the Presidency is therefore restricted to constituent peoples only.

The provisions outlined above were designed as a means to ensure balanced representation of the constituent peoples in order to sustain the fragile peace. Bardutzky states that '(...) the constitutional arrangements can be seen as a complex set of strong checks and balances between the three constituent peoples'.¹³ However, this focus on the rights of constituent peoples comes at the expense of 'others'. Edwards argues that the constitutional provisions '(...) have tended to reinforce the delicate politico-ethnic balance between the three main groups. Against this background, the rights of other ethnic or minority groups who were not central parties to the conflict have been ignored'.¹⁴ Of these ethnic or minority groups, the Roma minority is the largest¹⁵ and encounters the most difficulties.

B. The Roma Minority in BiH

Statistics prepared by the Council of Europe indicate that there are between 10 and 12 million Roma in Europe.¹⁶ They are Europe's largest minority group.¹⁷ Yet they are also subject to wide-spread prejudice and discrimination, leading to social,

⁹ *ibid* Article IV.

¹⁰ *ibid* Article III (c).

¹¹ *ibid* Article IV (1).

¹² *ibid* Article IV(2) which states: 'The House of Representatives shall comprise 42 Members, two-thirds elected from the territory of the Federation, one-third from the territory of the Republika Srpska.' However, Article IV (3) (b) does restrict the positions of Chair and Deputy Chair of the chamber only to constituent peoples.

¹³ Bardutzky (n 4) 310.

¹⁴ A Edwards, 'New Roma Rights Legislation in Bosnia and Herzegovina: Positive, Negative or Indifferent?' (2005) 9 (4) *IJHR* 465, 467.

¹⁵ V Tolic, 'Minorities in Bosnia and Herzegovina: Situation After Dayton' (2006)

<<http://www.migrationeducation.org/20.1.html?&rid=35&cHash=c2d15a5879d544551ee9d0c32697fd8>> accessed 31 October 2012.

¹⁶ Council of Europe Roma and Travellers Division, 'Statistics' updated 14 September 2010

<http://www.coe.int/t/dg3/romatravellers/default_en.asp> accessed 10 March 2013.

¹⁷ H O'Nions, *Minority Rights Protection in International Law: The Roma of Europe* (Ashgate 2007) 2.

political and economic exclusion.¹⁸ In BiH, this Europe-wide issue is exacerbated by historical tensions. Whilst Roma fought on all sides during the Bosnian War, they did not formally ally themselves with any of the combatant groups. Consequently, Roma have been left isolated in post-Dayton BiH, facing prejudice and often violence from all sides wherever they live.¹⁹ Although it is unclear exactly how many Roma currently reside in BiH,²⁰ most estimates place the number between 30,000 and 50,000.²¹ The provisions of the Constitution therefore serve to exclude a sizable minority that is in particular need of effective political inclusion from full participation in state government.

3. The Protection of Human Rights in BiH

A. The Constitution of BiH

The Constitution contains a number of provisions designed to ‘(...) ensure the highest levels of internationally recognized human rights and fundamental freedoms’.²² Annex 1 to the Constitution lists 15 additional human rights agreements applicable in BiH, including the 1966 International Covenant on Civil and Political Rights and its Optional Protocols, the 1965 International Convention on the Elimination of all Forms of Racial Discrimination, and the 1994 Framework Convention for the Protection of National Minorities. By virtue of their inclusion in the Constitution, these instruments are binding on the state.²³ Additionally, the

¹⁸ *ibid* chapter 1.

¹⁹ *ibid* 9.

²⁰ Gathering up-to-date information on the size of minority groups in war-torn BiH has proved extremely difficult (see the Advisory Committee on the Framework for National Minorities, ‘Opinion on Bosnia and Herzegovina’ ACFC/INF/OP/I (2005) 003 [52]. Roma representatives have expressed the view that a new census should be undertaken in the near future in order to accurately establish the number of Roma in the state (European Roma Rights Centre, ‘The Non-Constituents: Rights Deprivation of Roma in Post-Genocide Bosnia and Herzegovina’ (2004) Country Report Series, No 13, at 20 <<http://www.errc.org/cms/upload/media/00/28/m00000028.pdf>> accessed 25 November 2012).

²¹ Minority Rights Group <<http://www.minorityrights.org/?lid=2471>> accessed 10 November 2012. A 1991 census recorded only 9000 Roma, but this may be attributable to a reluctance to declare minority status at a time of great political upheaval. Roma NGOs in BiH have estimated that the number of Roma may be as high as 100,000, but this figure has not been widely accepted (Council of Roma of the Federation of Bosnia and Herzegovina, ‘Comments on the Implementation of the Framework Convention on the Protection of National Minorities in BiH’ <[http://www.nrc.ch/8025708F004CE90B/\(httpDocuments\)/E1DDFCE39A48E213802570B700587783/\\$file/Council+of+Roma+1.pdf](http://www.nrc.ch/8025708F004CE90B/(httpDocuments)/E1DDFCE39A48E213802570B700587783/$file/Council+of+Roma+1.pdf)> accessed 25 November 2012).

²² Constitution (n 1) Article II (1).

²³ Regarding specific instruments, see the Human Rights Committee, ‘Concluding Observations of the Human Rights Committee: Bosnia and Herzegovina’ CCPR/C/BIH/CO/1 22 November 2006 [4], and the Committee on the Elimination of Racial Discrimination, ‘Concluding Observations of the Committee on the Elimination of Racial Discrimination: Bosnia and Herzegovina’ CERD/C/BIH/CO/6, 11 April 2006 [6]. However, it should also be noted that commentators have expressed differing views as to the status of these instruments (see Z Pajic, ‘A Critical Appraisal of Human Rights Provisions of the Dayton Constitution of Bosnia and Herzegovina’ (1998) 20(1) Hum. Rts.Q. 125, 131 & Edwards (n 14) 467.

Constitution provides that the provisions of the European Convention on Human Rights (ECHR)²⁴ shall apply directly in Bosnia and Herzegovina, and shall have priority over all other law.²⁵ The rights and freedoms set out in both Article II and the annexed agreements are to be enjoyed without discrimination on any ground.²⁶ It has been suggested that '[t]here are few countries in which human rights are as richly guaranteed as they are in Bosnia and Herzegovina.'²⁷

One such human right ostensibly guaranteed to the people of BiH is the right to enjoy free elections without distinction on the basis of ethnicity. Participation in free elections is a means of giving voice to the political will of the people.²⁸ Free elections provide a platform for political participation, and a democratically elected government is an essential requirement for the wider protection of human rights within a state.²⁹ Where an ethnic group is discriminated against in this context, their ability to engage in effective political participation is restricted. Consequently, their interests are unlikely to be effectively represented. This creates an acute risk of social and economic exclusion. The right to enjoy free elections without distinction on the basis of ethnicity is thus fundamentally important. However, the next part of this article will demonstrate that there is a yawning gap between the right to free elections enshrined in text and the rights in practice afforded to the Roma minority by the Constitution of BiH.

B. International Human Rights Instruments

(i) The International Covenant on Civil and Political Rights ('ICCPR')

One of the primary international instruments for the protection of civil and political rights is the ICCPR. Article 25 states that:

Every citizen shall have the right and the opportunity, without any of the distinctions mentioned in article 2 and without unreasonable restrictions: (...) (2) To vote and to be elected at genuine periodic elections which shall be by universal and equal suffrage and shall be held by secret ballot, guaranteeing the free expression of the will of the electors.

Article 2 provides that states must recognise the Covenant rights 'without distinction of any kind'. Article 26 provides a further autonomous right prohibiting discrimination in law or in fact in any field regulated by public bodies, and

²⁴ The European Convention for the Protection of Human Rights and Fundamental Freedoms, as amended.

²⁵ Constitution (n 1) Art II (2).

²⁶ Constitution (n 1) Article 11 (4) states that: 'The enjoyment of the rights and freedoms provided for in this Article or in the international agreements listed in Annex I to this Constitution shall be secured to all persons in Bosnia and Herzegovina without discrimination on any ground such as sex, race, colour, language, religion, political or other opinion, national or social origin, association with a national minority, property, birth or other status.'

²⁷ ERRC Report (n 20) 9.

²⁸ United Nations Centre for Human Rights, 'Human Rights and Elections: A Handbook on the Legal, Technical and Human Rights Aspects of Elections' (1994) HR/P/PT/2 [2].

²⁹ *ibid* [4].

guarantees equal and effective protection from discrimination.³⁰ Any differentiation in treatment must be reasonable and objective, and must be in pursuit of a legitimate aim.³¹ These provisions secure the right and opportunity to be elected at genuine periodic elections for every citizen without any distinction,³² including distinction on the basis of race or ethnicity.³³ It is universally accepted that Roma are an ethnic minority,³⁴ and any distinction on the basis of their Roma status is thus a distinction on the basis of ethnicity.

Responsibility for monitoring the implementation of the ICCPR rests with the UN Human Rights Committee. This Committee has recognised the incompatibility of Articles IV and V of the Constitution with Articles 2, 25 and 26 of the ICCPR and called for constitutional reform '(...) with a view to adopting an electoral system that guarantees equal enjoyment of the rights under Article 25 of the Covenant to all citizens irrespective of ethnicity.'³⁵ It has called for these reforms to be made '(...) as a matter of urgency'.³⁶ The exclusion of 'others' from both institutions clearly removes from those who are defined as such the right and opportunity to be elected in accordance with Article 25. However, the incompatibility of the Constitution with Article 25 is also much wider: the exclusion of 'others' deprives the electorate as a whole of the right to 'free expression', should they wish to elect a Roma candidate.³⁷ Additionally, in the Federation the Bosniac and Croat delegates to the House of Peoples are selected by Bosniac and Croat members of the House of Peoples of the Federation respectively. Thus, by virtue of their ethnicity, Roma members are deprived of their right to participate in the selection of delegates. Articles 2, 25 and 26 are therefore violated on three separate fronts and BiH is clearly failing to meet its obligations under the ICCPR.

³⁰ UN Human Rights Committee, General Comment No 18, 'Non-discrimination' 10 November 1989 [14].

³¹ *ibid* [13].

³² *Ignatane v Latvia*, Communication No 884/1999, CCPR/C/72/D/884/1999 [7.3].

³³ Although Article 26 only refers specifically to race, race and ethnicity have been interpreted as related concepts. This is articulated by the ECtHR in *Sejdić* at [43], where the court takes the view that '[w]hereas the notion of race is rooted in the idea of biological classification of human beings into subspecies on the basis of morphological features such as skin colour or facial characteristics, ethnicity has its origin in the idea of societal groups marked in particular by common nationality, religious faith, shared language, or cultural and traditional origins and backgrounds. Discrimination on account of a person's ethnic origin is a form of racial discrimination'. This statement is of course made in the context of interpreting the non-discrimination provisions contained in the ECHR, but the court makes reference to the definition adopted by the International Convention on the Elimination of All Forms of Racial Discrimination and that adopted by the European Commission against Racism and Intolerance. There is clear international and regional consensus that ethnic discrimination falls within the ambit of racial discrimination, and it is on this basis that this article will proceed.

³⁴ O'Nions, *Minority Rights Protection in International Law: The Roma of Europe* (n 17) 185.

³⁵ UN Human Rights Committee (n 23) [8]. This is reiterated in the Committee's subsequent observations (UN Human Rights Committee, 'Concluding Observations on the Second Periodic Report of Bosnia and Herzegovina', adopted by the Committee at its 106th Session (15 October-2 November 2012) CCPR/C/BIH/CO/2, 13 November 2012 [6]), which were made in light of the *Sejdić* judgment.

³⁶ *ibid*.

³⁷ ERRC Report (n 20) [5.3].

(ii) The International Convention on the Elimination of all Forms of Racial Discrimination ('ICERD')

BiH is similarly failing to meet its obligations under the ICERD, and the Committee on the Elimination of Racial Discrimination (CERD) has made comments very similar to those of the Human Rights Committee noted above. The ICERD deals exclusively with racial discrimination, defined as:³⁸

Any distinction, exclusion, restriction or preference based on race, colour, descent, or national or ethnic origin which has the purpose or effect of nullifying or impairing the recognition, enjoyment or exercise, on an equal footing, of human rights and fundamental freedoms in the political, economic, social, cultural or any other field of public life.

Article 5(c) specifically guarantees equality in the enjoyment of political rights, '(...) in particular the right to participate in elections – to vote and to stand for election – on the basis of universal and equal suffrage, to take part in the Government as well as in the conduct of public affairs at any level and to have equal access to public service'.³⁹ In dealing with Articles IV and V of the Constitution, the Committee notes that '[l]egal distinctions that favour and grant special privileges and preferences to certain ethnic groups are not compatible with Articles 1 and 5(c) of the Convention'.⁴⁰ The CERD, like the Human Rights Committee, urges the State to amend the Constitution to ensure that all citizens have the right and opportunity to vote and stand for election without distinction on the grounds of ethnicity.⁴¹ The Committee adopts an interesting perspective, however, in that it addresses the issue as being one of additional rights extended to the constituent peoples on the basis of their ethnic affiliation, rather than as a matter of 'others' being subject to an exclusion or restriction of their rights.⁴² Regardless, the outcome remains the same, insofar as 'others' are unable to participate in elections on an equal footing and are thus subject to discrimination incompatible with the ICERD. This is a further failing on the part of BiH to fulfil its international obligations by ensuring that members of the Roma minority enjoy the right to participate in free elections without distinction on the basis of their ethnicity.

C. Regional Human Rights Instruments

(i) The Framework Convention for the Protection of National Minorities ('FCNM')

This international failure is duplicated on a European level. Article 4 of the Framework Convention for the Protection of National Minorities provides that:

³⁸ ICERD Article 1.

³⁹ *ibid* Article 5(c).

⁴⁰ Committee on the Elimination of Racial Discrimination (n 23) [11].

⁴¹ *ibid*. This is reiterated in the CERD's subsequent Observations in 2010 (Committee on the Elimination of Racial Discrimination, 'Concluding Observations of the Committee on the Elimination of Racial Discrimination: Bosnia and Herzegovina' CERD/C/BIH/CO/7-8, 23 September 2010 [7]).

⁴² *ibid*.

1. The Parties undertake to guarantee to persons belonging to national minorities the right of equality before the law and of equal protection of the law. In this respect, any discrimination based on belonging to a national minority shall be prohibited.
2. The Parties undertake to adopt, where necessary, adequate measures in order to promote, in all areas of economic, social, political and cultural life, full and effective equality between persons belonging to a national minority and those belonging to the majority. In this respect, they shall take due account of the specific conditions of the persons belonging to national minorities.

Article 15 obliges states to '(...) create the conditions necessary for the effective participation of persons belonging to national minorities in cultural, social and economic life and in public affairs, in particular those affecting them.'⁴³ BiH recognises Roma as 'the largest national minority.'⁴⁴ When taken together, these provisions create an obligation on states to take positive measures to ensure both equality and effective participation in public affairs for Roma. BiH is currently failing to meet these obligations. In their first Opinion on the implementation of the FCNM in BiH, the Advisory Committee concluded that the provisions of Articles IV and V of the Constitution were discriminatory, stating that:⁴⁵

While it may be said that they pursue a legitimate aim, namely to ensure equal representation of the three constituent peoples, their proportionality is questionable in terms of totally excluding in particular persons belonging to national minorities from accessing key-positions in public life. This therefore raises issues of compatibility with Article 4 of the Framework Convention.

These concerns were reiterated in the 2nd cycle, when the Advisory Committee urged the state to pursue constitutional reform '(...) with a view to eliminating discrimination against persons who do not belong to the constituent peoples and to enabling them to participate effectively in public affairs'.⁴⁶ The Committee also took the view that the constitutional provisions were incompatible with Article 15, as BiH was failing to take necessary steps to ensure the effective participation of minorities in public affairs.⁴⁷ The provisions of the FCNM not only prohibit exclusion, but also require states to adopt 'adequate measures' to promote equality. BiH is currently failing in both respects.

⁴³ FCNM Article 15.

⁴⁴ Bosnia and Herzegovina Council of Ministers, 'Report of Bosnia and Herzegovina on Legal and Other Measures on Implementation of the Principles Determined in the FCNM' ACFC/SR (2004) 001 17. In the absence of any definition of national minorities within the Framework, it is left to states to make their own decisions as to whether they have national minorities, and who these minorities are.

⁴⁵ Advisory Committee on the Framework Convention for National Minorities (n 20) [39].

⁴⁶ Advisory Committee on the Framework Convention for National Minorities, 'Second Opinion on Bosnia and Herzegovina' ACFC/OP/II(2008) 005 [69]. A third Opinion should be available in the near future, but given the lack of reform to date it is likely to reiterate the same concerns.

⁴⁷ *ibid* [199].

(ii) The European Convention for the Protection of Human Rights and Fundamental Freedoms ('ECHR')

In Europe, the ECHR also protects the right to free elections without discrimination. Article 3 of Protocol No 1 states:

The High Contracting Parties undertake to hold free elections at reasonable intervals by secret ballot, under conditions which will ensure the free expression of the opinion of the people in the choice of the legislature.

Article 14 prohibits discrimination in the enjoyment of the Convention rights⁴⁸ whilst Article 1 of Protocol No 12 provides a general non-discrimination clause protecting the enjoyment of 'any right set forth in law'⁴⁹ without discrimination. The compatibility of the Constitution with these provisions was considered in *Sejdić and Finci v Bosnia and Herzegovina*,⁵⁰ which was heard by the Grand Chamber of the European Court of Human Rights in December 2009. Mr Sejdić was a member of the Roma minority, whilst Mr Finci was Jewish. Both alleged that their ineligibility as 'others' to stand for election both to the House of Peoples (first claim) and to the Presidency (second claim) amounted to discrimination. In relation to their first claim, the claimants relied on Article 14 in conjunction with Article 3 of Protocol No 1. Their second claim was based on Article 1 of Protocol No 12.⁵¹ The Grand Chamber defined discrimination as '(...) treating differently, without an objective and reasonable justification, persons in similar situations' and reiterated that an objective and reasonable justification must be proportionate to a legitimate aim.⁵² Further, it stated that where a distinction is based on race or ethnicity, reasonable and objective justification must be interpreted strictly.⁵³ The Grand Chamber ultimately took the view that the discrimination against 'others' within the Constitution lacked objective and reasonable justification. Consequently, it found that there was a violation of both Article 14, taken in conjunction with Article 3 of Protocol No 1, in relation to the House of Peoples,⁵⁴ and Article 1 of Protocol No 12 in relation to the Presidency. *Sejdić* is a legally binding judgment which clearly demonstrates that Articles IV and V of the Constitution are discriminatory in nature when considered alongside the right to free elections. Claridge has noted that '[i]n spite of the delicate political situation in Post-Conflict Bosnia, it is difficult to conclude that the ECtHR could have reached any other verdict in relation to the direct, ethnic discrimination contained in

⁴⁸Article 14 ECHR states that: 'The enjoyment of the rights and freedoms set forth in [the] Convention shall be secured without discrimination on any ground such as sex, race, colour, language, religion, political or other opinion, national or social origin, association with a national minority, property, birth or other status.'

⁴⁹ Article 1 of Protocol No 12 to the ECHR.

⁵⁰ *Sejdić* (n 6).

⁵¹ This was, in itself, significant as it was the first time the ECtHR had had the opportunity to consider Article 1 of Protocol No 12.

⁵² *Sejdić* (n 6) [42].

⁵³ *ibid* [44]. The court referred to *DH and Others v The Czech Republic* (2007) 47 EHRR 59 [196].

⁵⁴ The court did not examine whether there was a breach of Article 3 of Protocol No 1 taken alone.

Bosnia's Constitution'.⁵⁵ Given the prominence afforded to the ECHR in the Constitution,⁵⁶ it is clear that BiH is failing to meet some of its most fundamental obligations as a state.

4. *Can the Discriminatory Constitutional Provisions be Justified?*

The need to end conflict and secure peace provides some explanation for the 'unique'⁵⁷ institutional arrangements enshrined in the Constitution. However it is necessary to examine whether and to what extent the consequent discriminatory impact of these constitutional provisions can be justified by the political tensions which still exist in BiH, and by the need to ensure continued peace. There appears to be consensus amongst the bodies tasked with implementing the instruments discussed above that this question should be answered in the negative. The Committee on the Elimination of Racial Discrimination has stated that the measures '(...) may have been justified, or even initially necessary to establish peace following the armed conflict within the territory of the State party' but that '(...) the exigency for which the special privileges and preferences were undertaken has abated.'⁵⁸ The Advisory Committee recognised that although the Constitution and the institutional framework it sets out have been 'instrumental in securing stability' in BiH, the continued exclusion of national minorities from the Presidency and the House of Peoples could not be justified in the long term.⁵⁹ In *Sejdić*, the ECtHR also noted that the measures were initially motivated by the need to restore peace. Whether this is a legitimate aim was not established, as the court took the view that the continued exclusion of 'others' was not proportionate to the aim of ensuring peace. In reaching this conclusion, the court pointed to the considerable positive progress achieved in BiH since the end of the conflict,⁶⁰ alternative power-sharing mechanisms suggested by the Venice Commission,⁶¹ and the commitment made by BiH on several occasions to amend electoral legislation to comply with the ECHR.⁶² In these circumstances, the Grand Chamber concluded that the discriminatory impact of the constitution could not be justified.

The dissenting judgment of Judge Bonello in *Sejdić* adopts a different view. Although Judge Bonello agreed, *prima facie*, that those who do not identify as

⁵⁵ L Claridge, 'Protocol 12 and *Sejdić and Finci v Bosnia and Herzegovina*: A Missed Opportunity?' 2011 EHRLR 82, 86.

⁵⁶ Article II (2). See above.

⁵⁷ Advisory Committee on the Framework Convention for National Minorities, 'Second Opinion on Bosnia and Herzegovina' (n 46) [12].

⁵⁸ Committee on the Elimination of Racial Discrimination, 'Concluding Observations of the Committee on the Elimination of Racial Discrimination: Bosnia and Herzegovina' (n 23) [11].

⁵⁹ Advisory Committee on the Framework Convention for National Minorities, 'Second Opinion on Bosnia and Herzegovina' (n 46) [39].

⁶⁰ *Sejdić* (n 6) [47].

⁶¹ *ibid* [48], referring to the Venice Commission, 'Opinion on the Constitutional Situation in Bosnia and Herzegovina and the Powers of the High Representative' CDL-AD (2005) 004.

⁶² *ibid* [49]. The commitments referred to originate from those undertakings made by BiH in 2002 on accession to the Council of Europe, and are set out in Council of Europe Parliamentary Assembly Opinion No. 234 (2002) 15.

constituent peoples had been discriminated against by being denied the right to vote or to stand for election, he was not prepared to find that there had been a violation of the ECHR.⁶³ He opined that the majority ruling fails to take adequate account of the circumstances in which the Constitution was drafted, namely the pressing need to secure peace. Against this background, he accepts that the constitution may not be perfect. He stresses, however, that it embodies a carefully crafted set of checks and balances between the three constituent peoples designed to achieve and maintain a fragile peace in BiH.⁶⁴ He therefore questions whether it is appropriate for the Court to intervene in the circumstances.⁶⁵ His argument is compelling: bar a small number of core rights, human rights must be exercised in conformity with the rights of others and with due attention to 'the over-riding social good.'⁶⁶ The right to free elections is not absolute, and ECtHR case-law recognises a wide range of circumstances in which the restriction of electoral rights is justified.⁶⁷ In *Sejdić*, Judge Bonello was of the view that requiring the Constitution to be amended so as to remove the discriminatory provisions would pose '(...) a clear and present danger of destabilising the national equilibrium'⁶⁸ and that such a risk justifies the restrictions imposed on the rights of Roma to vote in and stand for election. On these grounds, he found that there had been no violation of the ECHR. He argued that it is for the national authorities of BiH, not the ECtHR, to decide if and when constitutional reform is appropriate.⁶⁹

This dissenting judgment is out of step with the views of the relevant monitoring bodies discussed above, and with that of the majority in the Grand Chamber. It is submitted here that, although the political and historical background against which the Constitution was drafted may initially have justified provisions which had the unfortunate consequence of discriminating against minority groups by restricting their right to participate in free elections, such justification no longer exists. Furthermore, the existence of alternative measures for ensuring peace which do not discriminate against minority groups indicates that the current constitutional arrangements are not proportionate to the aim they purport to pursue. However this article will now demonstrate that achieving the necessary constitutional reform has proved problematic. It may be that Judge Bonello's more cautious approach foreshadows these difficulties, and is ultimately a more realistic assessment of the current political situation in BiH.

5. The Limits of Constitutional Reform

The exclusion of Roma from the highest levels of political participation in Bosnia and Herzegovina has had an extremely detrimental impact. Roma are subject to regular abuse of their civil, political, economic and social rights as a consequence of their

⁶³ *ibid* [52].

⁶⁴ *ibid*.

⁶⁵ *ibid* [52-53].

⁶⁶ *ibid* [53].

⁶⁷ *ibid* [54-55].

⁶⁸ *ibid* [55].

⁶⁹ *ibid* [54].

status as second-class, non-constituent peoples.⁷⁰ As discussed, Roma across Europe experience daily discrimination. In BiH, however, this discrimination is constitutionally-mandated and consequently significantly more difficult to combat. The powerful position of the constituent peoples allows persons so-defined to legislate in a way that protects their own interests. Whilst this is unsurprising given the history of political tension in BiH, the result is that the interests of minorities are sidelined. Edwards states that, in its present form, the Constitution '(...) marginalises persons not belonging to one of the three main groups and reduces any chance for their meaningful political participation.'⁷¹ This particularly affects the substantial Roma minority. The Advisory Committee summed up the situation, stating that:⁷²

Although the Roma are by far the largest national minority group and the group facing the most serious difficulties, their numerical importance compared to other national minorities is not taken into due account in terms of representation in public affairs. This lack of adequate representation, coupled with social exclusion, results in very limited opportunities for them to participate effectively in public life.

There can be little doubt that exclusion from the House of Peoples and the Presidency therefore reinforces the vulnerability of the Roma,⁷³ a group already subject to prejudice and poverty in post-war BiH. Roma struggle to access key services including housing, healthcare and education.⁷⁴ In its 2012 World Report, Human Rights Watch noted a 99% unemployment rate for members of the Roma community.⁷⁵ Constitutionally-mandated discrimination thus has a highly detrimental impact on the daily lives of Roma in BiH.

Constitutional reform is the 'first and most necessary'⁷⁶ step to ending this discrimination and the detrimental impact it has on the Roma community. As well as creating the possibility for Roma to participate in the highest levels of state government, such reform would send a powerful message that continued discrimination against Roma at any level of state, entity, or local government will not be tolerated. Yet constitutional reform alone cannot fully address the widespread discrimination experienced by the Roma minority. Certain models of constitutional reform that would eliminate discrimination - for example, providing that delegates to the House of Peoples be appointed by the Council of Ministers - would not in fact do anything to enfranchise minority groups and provide them with a means of

⁷⁰ ERRC Report (n 20) 10. A detailed examination of these wider issues is out with the scope of this article.

⁷¹ Edwards (n 14) 475.

⁷² Advisory Committee on the FCNM (n 45) [200].

⁷³ ERRC Report (n 20) 12; Toliv (n 15).

⁷⁴ Human Rights Watch, 'Second Class Citizens: Discrimination Against Roma, Jews, and Other National Minorities in Bosnia and Herzegovina' 2012, <<http://www.hrw.org/reports/2012/04/04/second-class-citizens-0>> accessed 23 March 2013.

⁷⁵ Human Rights Watch, 'World Report 2012: Bosnia and Herzegovina' 2012. <<http://www.hrw.org/world-report-2012/bosnia-and-herzegovina>> accessed 20 March 2013.

⁷⁶ Human Rights Watch, 'Second Class Citizens: Discrimination against Roma, Jews, and Other National Minorities in Bosnia and Herzegovina' (n 74) 7.

expressing their political will.⁷⁷ Even if enfranchised by reform, there remains a real risk that political culture will continue to prevent minority groups from actively participating in the highest levels of state government.⁷⁸ Furthermore, piecemeal constitutional reform carried out in haste may exacerbate unresolved ethnic tensions within BiH, and may in turn cause the state to disintegrate along ethnic lines.⁷⁹ This is an outcome that should be avoided at all costs.

Whilst constitutional reform is a necessary condition for ending discrimination against Roma in BiH, it is a task that must be approached with extreme caution. Constitutional reform is unlikely to fully address the deep-rooted prejudice and discrimination experienced by the Roma minority in BiH. Such reform, therefore, must take place as part of a broader programme of reform at state, entity and local level, designed to challenge the current political culture which excludes Roma participation and to address the deep-rooted prejudice against the Roma community in BiH.

6. The Progress of Constitutional Reform

The decision in *Sejdić* has the potential to form the impetus for constitutional reform designed to bring BiH into line with its obligations under the regional and international instruments annexed to the Constitution. Yet Bardutzky sounds a note of caution when he states that '[t]he implications of this judgement for the constitutional development of BiH are perhaps as vast as they are unpredictable (...) let us believe that the judgement will serve as an impulse for constitutional reform rather than dissolution or diminished functionality of the State.'⁸⁰ More than three years have now passed since the Grand Chamber's judgment and such constitutional reform has not been forthcoming.

In 2010, the Council of Europe expressed its serious concern regarding this lack of progress, noting that two initiatives aimed at implementing constitutional reform had failed to achieve any positive results.⁸¹ A subsequent report⁸² highlighted that although BiH had adopted an Action Plan for implementation of the *Sejdić* judgment in February 2010, and formed a working group to address the issue, key stakeholders in the process had not attempted to negotiate proposals capable of achieving consensus. The report considered the working group a 'missed opportunity' to achieve constitutional reform.⁸³ Consequently, the General Elections

⁷⁷ International Crisis Group, 'Bosnia's Gordian Knot: Constitutional Reform' July 2012, Europe Briefing No.68 [6] <<http://www.crisisgroup.org/en/regions/europe/balkans/bosnia-herzegovina/b068-bosnias-gordian-knot-constitutional-reform.aspx>> accessed 12 March 2013. This would, however, be very much contrary to the spirit of the *Sejdić* judgment.

⁷⁸ *ibid* [1].

⁷⁹ *ibid*.

⁸⁰ Bardutzky (n 4) 332.

⁸¹ The so-called 'Pud process' and 'Butmir process'. See Council of Europe Parliamentary Assembly Resolution 1701 (2010) [4-5].

⁸² Committee on the Honouring of Obligations and Commitments by Member States of the Council of Europe (Monitoring Committee), 'The Urgent Need for a Constitutional Reform in BiH' Doc 12222, 27 April 2010.

⁸³ *ibid* [18].

in October 2010 were once again held in accordance with constitutional provisions contrary to the ECHR and other regional and international instruments.

Since the 2010 elections, there has been significant political upheaval in BiH. These elections resulted in a coalition government, which did not take office until February 2012. The coalition collapsed on 31st May 2012 and the outcome of continued reshuffling remains uncertain.⁸⁴ Unsurprisingly, '(...) the tensions linked to the reshuffling of the governments at State level and in the Federation have not been conducive to progress on the implementation of the *Sejdić-Finci* ruling'⁸⁵ and further attempts to bring forward proposals for constitutional amendment have proved unsuccessful. In October 2011, an Interim Joint Parliamentary Committee was established with the specific task of drafting constitutional and electoral reforms that would give effect to the *Sejdić* ruling. On 1st December 2011, that Committee officially announced its failure to draft such reforms.⁸⁶ In June 2012, representatives from the European Commission and BiH's main political parties agreed a 'Road Map' for BiH's application for membership of the European Union, by which proposals for constitutional amendment would be brought forward by 31st August 2012.⁸⁷ This deadline was not met.⁸⁸ Consequently, in October 2012 the European Commission noted that '(...) there has been little progress in compliance with the ECtHR judgement in the *Sejdić-Finci* case'⁸⁹ and that '[i]t remains essential to implement the ECtHR judgment in the *Sejdić-Finci* case in order to comply with the ECHR.'⁹⁰ To date, no credible proposals for reform have been brought forward by any group tasked with doing so.

It appears, therefore, that attempts to negotiate reform may have reached an impasse. Whilst the necessity of reform is accepted by the authorities in BiH,⁹¹ deep-seated historical tensions mean that the political representatives of the constituent peoples charged with negotiating reform are concerned primarily with preserving the rights of their own people.⁹² Consequently, negotiations to date have focused on models for reform that would implement the *Sejdić* judgment but that would also reinforce the dominant position of the main political parties' leadership.⁹³ Minority groups, including representatives of the Roma minority, have been afforded very

⁸⁴ European Commission, 'Bosnia and Herzegovina 2012 Progress Report' SWD (2012) 335, 10 October 2012 [2.1].

⁸⁵ *ibid* [1.2].

⁸⁶ Council of Europe Parliamentary Assembly Resolution 1855 (2012) [4].

⁸⁷ European Commission, 'Joint Conclusions from the High Level Dialogue on the Accession Process with Bosnia and Herzegovina and the Road Map for BiH's EU membership application' 27 June 2012 MEMO/12/503.

⁸⁸ European Commission, 'Bosnia and Herzegovina 2012 Progress Report' (n 84) [2.1].

⁸⁹ *ibid*.

⁹⁰ *ibid* [2.2].

⁹¹ International Crisis Group, 'Bosnia's Gordian Knot: Constitutional Reform' (n 77) 8. See also the comments made by BiH's Foreign Minister Zlatko Lagumdžija in April 2012 (Council of Europe Press Release, "'No back-sliding" in move to European standards says Foreign Minister of Bosnia and Herzegovina' 25 April 2012.

<http://www.assembly.coe.int/ASP/NewsManager/EMB_NewsManagerView.asp?ID=7645> accessed 12 March 2013.

⁹² Council of Europe Press Release, 'Commissioner Füle and Secretary General Jagland regret the lack of progress in implementing the *Sejdić-Finci* judgement' DC043 (2013), 8 April 2013.

⁹³ International Crisis Group, 'Bosnia's Gordian Knot: Constitutional Reform' (n 77) 1.

little opportunity to participate in attempts to negotiate reform.⁹⁴ Without a genuine willingness to compromise on the part of those involved in the negotiations, it is difficult to see how attempts to bring about reform can progress further.

The Council of Europe *Ad Hoc* Committee's Observation Report on the 2010 Elections states that '(...) the Parliamentary assembly should not allow the next general elections in BiH to take place under conditions incompatible with the ECHR.'⁹⁵ In January 2012, the Council of Europe warned that '[i]f the necessary amendments are not adopted in good time before the next elections in 2014, the continued membership of Bosnia-Herzegovina in the Council of Europe may be at stake.'⁹⁶ Further, the Council of the European Union has made it very clear that 'a credible effort' by BiH to bring its constitution into compliance with the ECHR by implementing the *Sejdić* judgment is necessary before a Stabilisation and Association Agreement (SAA) between the EU and BiH signed in 2008 can enter into force.⁹⁷ The Council has stated that '[a] satisfactory track record in implementing obligations under the SAA would be a key element for a credible membership application to be considered by the EU.'⁹⁸ As recently as April 11th 2013, the EU's Commissioner for Enlargement and European Neighbourhood Policy, Štefan Füle, reiterated that BiH must amend its Constitution to comply with the *Sejdić* ruling in order to move forward in the European integration process.⁹⁹ The continued presence of discriminatory provisions in the Constitution is therefore directly contributing to delays in BiH's progress towards EU membership. He also echoed comments made by Secretary-General Jagland of the Council of Europe that '[a]nother election violating the European Convention on Human Rights would be unacceptable, seriously undermining the legitimacy and credibility of the country's elected bodies'.¹⁰⁰ There is thus a very real risk that, should constitutional reform not precede the 2014 general elections, the resulting elected bodies will not be recognised by either the Council of Europe or the Institutions of the European Union. Whilst neighbouring states progress towards European Union membership, BiH risks finding itself increasingly politically isolated as a result of its failure to take concrete steps towards constitutional reform.

It is clear that BiH will be subject to continued and increasing external pressure to make the necessary reforms and fulfil its obligation to allow Roma participation in free elections without discrimination. The Council of the European Union has recently decided to extend its option of imposing sanctions against obstructive politicians in BiH, providing the Council with a means of exerting pressure on those political leaders who fail to prioritise the EU agenda over political

⁹⁴ Minority Rights Group, 'Minorities Side-lined in Panel to Oversee Change' 3 November 2011 <<http://www.minorityrights.org/11078/press-releases/minorities-sidelined-in-panel-to-oversee-change.html>> accessed 12 March 2013.

⁹⁵ Council of Europe Bureau of the Assembly, 'Observation of the General Elections in BiH (3rd October 2010)' Doc 12432 11 November 2010 [44].

⁹⁶ Council of Europe Parliamentary Assembly Resolution 1855 (2012) [5].

⁹⁷ Council of the European Union, 'Council conclusions on Bosnia and Herzegovina' 3179th Foreign Affairs Council meeting Luxembourg, 25 June 2012 [3].

⁹⁸ Council of the European Union, 'Council conclusions on Bosnia and Herzegovina' 3076th Foreign Affairs Council meeting, Luxembourg, 21 March 2011 [3].

⁹⁹ European Commission, 'Statement by Commissioner Štefan Füle (n 3).

¹⁰⁰ *ibid.*

and ethnic interests.¹⁰¹ Yet the delays to date, and the repeated failure to reach agreement on proposed reform, indicate that the political will to identify a workable solution is lacking. Whether further discussions and the threat of sanction can address this issue remains to be seen. Until these difficulties can be overcome, BiH's progress towards becoming '(...) a modern, Euro-compatible and functional democracy in which every citizen, regardless of his or her ethnic affiliation, enjoys the same rights and freedoms'¹⁰² has stalled and its membership of the Council of Europe is at risk. This is hugely concerning.

7. Conclusion

There can be no doubt that Articles IV and V of the Constitution of BiH currently discriminate against Roma and prevent them from enjoying the right to participate in free elections on an equal footing with those who enjoy the status of constituent peoples. BiH is thus failing to fulfil its obligations under the ICCPR and ICERD on an international level, and the ECHR and FCNM on a regional level. This has the result of excluding Roma from political participation at the highest levels, reinforcing their position as an isolated, alienated and disadvantaged minority group. The comments of international and regional monitoring bodies make clear that urgent constitutional reform is necessary to bring BiH into line with its obligations, and ultimately to end the detrimental exclusion of Roma from standing and voting in free elections for the highest state institutions. Such reform is a necessary, albeit not sufficient, condition for ending the daily discrimination and prejudice experienced by the Roma minority in BiH. The lengthy delay in achieving reform is extremely concerning, and indicates that the issue goes to the root of the political tensions which continue to exist in BiH. The state now finds itself subject to increasing external pressure from the monitoring bodies, the Council of Europe and the institutions of the EU to bring forward credible proposals for reform as a matter of urgency. These calls for reform have also been reiterated on an international level, notably in a 2010 speech by the then US Secretary of State Hillary Clinton.¹⁰³ One might hope that the swift approach of the 2014 general elections may provide the impetus for reform within the next year. Yet there remains unwillingness among the political leaders of BiH to compromise in order to reach agreement on proposed reform. Unless this impasse can be broken, there is a real risk that 2013 may be another year of inactivity. Until such time as reform takes place, the policy of BiH towards the Roma minority is clearly discriminatory and far from satisfactory. It is also a very real obstacle in BiH's progress from war-torn state to modern, European, democracy.

¹⁰¹ A Alic, 'Europe prolongs threat of Bosnian sanctions' *Southeast Europe Times* (25 March 2013) <http://www.setimes.com/cocoon/setimes/xhtml/en_GB/features/setimes/features/2013/03/25/feature-03> accessed 28 April 2013.

¹⁰² Council of Europe Parliamentary Assembly Resolution 1855 (2012) [6.1].

¹⁰³ Associated Press, 'Hillary Clinton urges Bosnia to make EU membership a priority' *The Guardian* (London 12 October 2010) <<http://www.guardian.co.uk/world/2010/oct/12/hillary-clinton-bosnia-european-union>> accessed 28 April 2013.

The Utility of the NESS Test of Factual Causation in Scots Law

EUAN WEST*

Abstract

It is well accepted in Scots law that, in order for a delictual claim to be successful, the pursuer must establish factual causation. In most cases, he or she will have to satisfy the 'but for' test, however because this can lead to harsh results the courts have sometimes adopted a more relaxed approach to causation. Thus, there are various exceptions to the general rule (namely the 'but for' test) including the 'material contribution' test adopted in Wardlaw v Bonnington Castings Ltd.¹ This article seeks to challenge the current Scots law approach to causation and to consider the merits of replacing the 'but for' test, along with the various exceptions thereto, with a single test of factual causation known as the NESS test. It will be contended that there are three main 'criteria' against which the utility of a causation test can be gauged and, with close reference to these 'criteria', it will be argued that the current Scots law approach to factual causation is inadequate. This article concludes that NESS satisfies all three 'criteria' and hence constitutes a promising alternative to the current law; one that the Scottish courts would do well to consider.

1. Introduction

In a 1985 article,² Richard Wright propounded the NESS test of factual causation. Building on the work of Hart & Honoré, who had developed a largely identical test,³ and Mackie,⁴ Wright argued that a cause in fact should be defined as a necessary element in a sufficient set (hereinafter a 'NESS'). Since then, NESS has become a very significant issue in American academic spheres, with even the test's most ardent critics calling it the '(...) new supplement to the but-for test for the twenty-first century.'⁵ It has also been said that scholarship surrounding NESS is the 'most

* Graduate, School of Law, University of Aberdeen.

¹ *Wardlaw v Bonnington Castings Ltd* 1956 SC (HL) 26 (hereinafter '*Wardlaw*').

² R Wright, 'Causation in Tort Law' (1985) 73 *California Law Review* 1735, 1788-1803.

³ HLA Hart & AM Honoré, *Causation in the Law* (1st edn, Oxford University Press 1959) 104-108, 116-119, 216-229 (as cited in J Stapleton, 'Choosing what we mean by "causation" in the law' (2008) 73 *Missouri Law Review* 433, 459).

⁴ JL Mackie, *The Cement of the Universe: A Study of Causation* (Clarendon Press 1974) chapter 3 (as cited in R Wright, 'The NESS Account of Natural Causation: A Response to Criticisms,' in R Goldberg (ed), *Perspectives on Causation* (Hart Publishing 2011) 288).

⁵ DA Fischer, 'Insufficient Causes' (2005-2006) 94 *Kentucky Law Journal* 277.

successful influential work in this area [tort/delict]'.⁶ Although it does not appear that NESS is currently used in American judicial practice, its prominence in American legal discourse cannot be denied.⁷

Such enthusiasm is not shared in Scotland. Although there are some UK academics who champion NESS,⁸ for the most part, the test rarely features in academic writings on delict. Indeed, it may be fair to say that, outwith American academic circles, there is a general indifference towards the test. As Hogg recently remarked, an internet search of 'NESS' will yield plenty of results concerning Loch Ness but hardly any relating to its causal namesake.⁹ Such apathy may, however, be unwarranted. In recent years, Scotland, and indeed the UK as a whole, has struggled with causation. The standard 'but for' test leads to unfair results and the various exceptions introduced to redress the harshness of the general rule, for instance the notorious '*Fairchild*' exception,¹⁰ have rendered the law of causation unprincipled and chaotic. Herein lies the potential appeal of NESS.

With the difficulties with the current law on causation in mind, the purpose of this article is to investigate the potential utility of NESS in Scots law. Part II will demonstrate why the current state of factual causation is unsatisfactory. Part III will then assess whether NESS is a rational test and, although the issues raised here may seem theoretical, they will all bear heavily upon the question of NESS' practical benefits and drawbacks. Part IV will focus on the practical utility of NESS. Throughout, but especially in Parts II and IV, reference will be made to various criteria for an ideal test of factual causation. By assessing the extent to which NESS complies with these criteria, it will be possible to evaluate the test's practical utility.

It is important to underline that NESS is a test of *factual* causation. Typically, once the pursuer in a delictual claim has proven that the defender was in breach of a duty of care, he or she will have to establish that the defender's wrongdoing was a factual cause of his or her loss.¹¹ This is an empirical enquiry, divorced from normative considerations. As such, it must not be confused with legal causation, an enquiry which does not really concern causation at all but rather issues such as whether the defender, despite having factually caused the pursuer's injury, should

⁶ R Fumerton and K Kress, 'Causation and the Law: Preemption, Lawful Sufficiency and Causal Sufficiency' (2001) 64:4 Law and Contemporary Problems 83, 83.

⁷ In his most recent article regarding the test, Wright refers to various commentators, including the American Law Institute and several academics, who have attested to the merits of the NESS test. However, he does not cite any examples of the NESS test being expressly employed by the courts. See Wright, 'The NESS Account of Natural Causation: A Response to Criticisms' (n 4) 285.

⁸ Most notably Martin Hogg and Chris Miller. See especially: M Hogg, 'Developing Causal Doctrine' in Goldberg (n 4) & Chris Miller, 'NESS for Beginners' in Goldberg (n 4).

⁹ *ibid* (Hogg) 43.

¹⁰ *Fairchild v Glenhaven Funeral Services Ltd* [2003] 1 AC 32 (hereinafter '*Fairchild*'). For a detailed discussion of this case, see part II below. Note that, although this was an English appeal to the House of Lords, it is all too relevant to the Scots law of causation. This is because, unlike other areas of law, the Scots law of delict and the English law of torts have much in common. A famous example of this is the case of *Donoghue v Stevenson* [1932] AC 562, a landmark decision for both Scots law and English law relating to another area of delict: duty of care. Note further that, in *Fairchild*, reference was made to *McGhee v National Coal Board* [1973] 1 WLR 1 (hereinafter '*McGhee*') and *Wardlaw* (n 1), both of which were Scottish appeals to the House of Lords. Thus, as far as causation in delict/tort is concerned, there is much overlap between English and Scots law.

¹¹ J Thomson, *Delictual Liability* (Tottel Publishing Ltd 2009) 139.

nevertheless not be held liable for unforeseeable consequences of the breach.¹² Therefore, for the avoidance of ambiguity, all subsequent references to ‘causation’ will denote the non-normative concept of factual causation. Although some commentators,¹³ are of the view that causation cannot be separated from normative considerations, and that its meaning varies according to legal context, it will be assumed, for present purposes, that some rationale exists for having a single, non-normative test of factual causation; that the factual-legal causation dichotomy is preferable to Lord Hoffmann’s proposed alternative. The purpose of this article is not simply to assess NESS in the void, but its utility as a *test of factual causation*.

2. Does Scots Law Need a New Test of Factual Causation?

In assessing the adequacy of Scots law’s current approach to factual causation and the possible need for reform, the following issue demands careful examination: what makes a ‘good’ test of causation? A search for the relevant criteria invites consideration of a further issue, namely the purpose of causation. This part will therefore begin by considering the purpose of causation, then, once it has been established *why* Scots law requires causation as a prerequisite to a successful delictual claim, it will be possible to infer various criteria against which to evaluate the current law.

A. The ‘Three Criteria’

In Schroeder’s view,¹⁴ there is no justifiable reason for the causation enquiry; no principled basis upon which to hold that, if both X and Y are negligent, but only Y’s wrong *causes* Z loss, Y alone is liable. In all normatively relevant respects, X and Y are in the same position: both are in breach of a duty of care owed to Z. However, because of something as ‘fortuitous’ as causation, they are treated differently.¹⁵ In short, as a basis for deciding whether a delictual claim should succeed or fail, causation has about as much moral legitimacy as the roll of a dice.

This is not to suggest, however, that causation serves no purpose. Certainly, the doctrine seems to lack a sound, principled basis but it could still be rationalised as a matter of policy. Arguably, the requirement to establish causation derives from a ‘floodgates’ policy whereby the law attempts to limit the category of persons who can be held liable to the pursuer.¹⁶ This is not to be confused with doctrines such as

¹² *Sayers v Harlow Urban District Council* [1958] 2 All ER 342, [1958] 1 WLR 623 CA.

¹³ See especially, Rt Hon Lord Hoffmann, ‘Causation’ in Goldberg (n 4) 9.

¹⁴ CH Schroeder, ‘Causation, Compensation and Moral Responsibility’ in DG Owen (ed) *Philosophical Foundations of Tort Law* (Clarendon Press 1995) 349.

¹⁵ *ibid.*

¹⁶ This sort of ‘floodgates’ policy has featured in other areas of the law of delict/tort. Consider, for example, the case of *Alcock v Chief Constable of South Yorkshire Police* 1992 1 AC 310. In that case, the House of Lords imposed various limitations on the category of persons who could claim for psychiatric injury suffered as a result of witnessing the Hillsborough disaster. For example, in order for a secondary victim to claim for psychiatric injury, it would be necessary for that person to establish that he or she had been present at the accident or its immediate aftermath. These rules were used by

duty of care, which limit the scope of the defender's liability by reference to normative considerations. Factual causation does not impede an otherwise successful claim on the basis of normative considerations. Rather, if only for expediency's sake, the law has seen it fit to 'draw a line' after which one cannot be liable. To do this using causation may be morally arbitrary but, since causation is an empirical fact the existence or non-existence of which is easily verified, the line is at least well-defined. The division between those who have a claim and those who do not may be arbitrary but it is also clear-cut.

When the purpose of factual causation is viewed in this way, various criteria for a suitable test can be inferred. For a start, if causation is meant to limit liability then an effective test will represent a meaningful bar to certain claims for, if causation becomes too easy to establish, the very purpose of the causal enquiry will be undermined. For ease of reference, this criterion will be termed 'limiting liability'.

In addition, since a clear line must be drawn between those defenders who caused the pursuer harm and those who did not, the causation test must guarantee consistency in the law and that like cases are treated alike. Theoretically, if multiple judges apply the test to a particular set of facts, they should all reach the same conclusion. To that end, the causation test should be simple, objective and unambiguous. If judges can depart from the usual test on grounds of fairness, inconsistencies will abound and the clear line that the causation enquiry strives to draw between claimants will be blurred. This requirement will be referred to in this article as 'clarity and consistency'.

It is proposed here that the third criterion for a useful causation test is fairness. Of course, this requirement is not absolute because, as explained above, the primary function of causation is to limit the scope of liability in a clear-cut way, even when that results in injustices. However, insofar as this primary function allows, the test must not be unduly harsh on the pursuer; it must not be so stringent as to never establish causation, barring liability in nigh every case. In this way, the test should recognise a fairly broad category of causation.

B. Does the 'But For' Test Satisfy the 'Criteria'?

In order to establish factual causation in Scots law, the pursuer must demonstrate that there is a link between the defender's breach of duty and the injury sustained. In other words, the court must be satisfied that, 'but for' the defender's wrongdoing, the pursuer's injury would not have occurred.¹⁷ In most cases, the 'but for' test is

the House of Lords to limit the extent of the defendants' liability and one of the policies underlying them was the notion that, if liability were not restricted, the 'floodgates' of litigation might be opened (see, for instance, Nolan LJ's comments at [383]). Although *Alcock* did not concern causation, but rather the requirement to establish a duty of care, the case is a telling illustration of how arbitrary limits on the scope of liability are often underpinned by 'floodgates' considerations. It is submitted here that the same sort of 'floodgates' policy invoked in *Alcock* may also account for the requirement to prove causation; a requirement which is, like the rules in *Alcock*, an arbitrary, non-normative restriction on liability.

¹⁷ A classic example of this test being applied is the case of *McWilliams v Sir William Arrol* [1962] 1 WLR 295, 1962 SC (HL) 70, in which it was held that the defenders' failure to provide a steel erector

perfectly capable of meeting the law's needs. For a start, it is fairly strict. In demanding causation as a matter of necessity, the test represents a decisive obstacle to many claims. The 'but for' test also offers the benefits of objectivity and clarity. Whether X's breach was necessary for Y's loss is often a fairly clear-cut issue and one on which judges seldom differ. Decisions such as *Kay's Tutor v Ayrshire and Arran Health Board*¹⁸ may appear harsh at first sight,¹⁹ but it must be remembered that causation's primary purpose is not to achieve fair results but certain ones; results that limit the scope of liability in a clear, effective manner. Applied consistently to each case, the 'but for' test serves this purpose well.

Nevertheless, insofar as this does not compromise the 'limiting liability' criterion, a causation test should still be fair, and it is in this regard that the 'but for' test is unsatisfactory. If, for example, two hunters (H1 and H2) simultaneously shoot a hill-walker,²⁰ H1's gunshot cannot be deemed necessary for the walker's death because H2 would have killed him anyway and, since H2's gunshot cannot be deemed necessary either, the 'but for' test holds that neither gunshot was a cause.

The test also struggles with causal indeterminacy, a problem that manifested itself clearly in *Fairchild*.²¹ In this well-known case, a man had been exposed to asbestos by various employers, as a result of which he contracted mesothelioma. However, owing to gaps in scientific knowledge, the House of Lords had no way of ascertaining which exposure had initiated the disease and so could not hold either one to be a 'but for' cause. This result seemed especially unfair considering that one of the exposures *must* have caused the cancer, it was simply not clear which. In sum, it is argued here that because the 'but for' test leads to such unjust results, it falls short of one of the key 'criteria' of a 'good' causation test: fairness.

C. Do the Exceptions to the 'But For' Test Satisfy the Criteria?

One alternative to the 'but for' test is the 'material contribution' rule, developed in *Wardlaw*, a Scottish case appealed to the House of Lords. That case concerned a worker who contracted pneumoconiosis by inhaling silica dust. Although the defender was in breach of a duty as regards some of the dust, he owed no such duty in respect of the rest. Because the pneumoconiosis could be attributed to multiple sources, the 'negligent' dust was not a 'but for' cause. Nevertheless, the House of

with a safety belt was not a 'but for' cause of him falling to his death because, even if he had been furnished with such a belt, he would not have worn it.

¹⁸ *Kay's Tutor v Ayrshire and Arran Health Board* 1987 SC (HL) 145 (hereinafter '*Kay's Tutor*').

¹⁹ The pursuer's son had suffered loss, namely deafness, and the hospital treating him had, in administering him with an overdose of penicillin, committed a breach of duty. Thus, in most normatively relevant respects, the claim in *Kay* seemed intuitively valid. It simply failed because causation could not be established, there being no evidence that a penicillin overdose could cause deafness.

²⁰ Stapleton (n 3) 442. The 'two hunters' example referred to by Stapleton is loosely based on the facts of *Cook v Lewis* [1951] SCR 830 (SCC) and *Summers v Tice* (1948) 33 Cal 2D 80 (as cited in E Adams, 'The Flexibility of Description and NESS Causation' (2010) 10 Journal Of Philosophy, Science and Law <http://www6.miami.edu/ethics/jpsl/archives/all/NESS_causation.html> accessed 8 June 2013. See footnote 14 of that article).

²¹ *Fairchild* (n 10).

Lords established that there was causation on the basis that the 'negligent' dust had materially contributed to the disease.

Unfortunately, this notion of 'material contribution' is so vague,²² so circular,²³ that a judge can effectively establish causation on the basis of intuition. The decision that factor X made a 'material contribution' is not based on some empirical quality that objectively identifies X as a cause; rather the judge has intuited that X *should* be deemed causal. Indeed, the normative connotations of a contribution that was 'material' or not '*de minimis*'²⁴ suggest that the doctrine would be better suited to legal causation. It may achieve fair results but, thanks to its vagueness, the rule fails to guarantee clarity and consistency in the law.

In addition to the *Wardlaw* rule, the House of Lords has supplemented the 'but for' test with an 'increase in risk' approach. This has its origins in *McGhee*,²⁵ another Scottish case in which a man claimed that his employer had caused him to contract dermatitis. The pursuer worked in a hot, dusty environment and, because of the absence of on-site washing facilities, cycled home unwashed. Although the defender had a duty to provide washing facilities, the court was not satisfied that he was in breach of any duty as regards the pursuer's hot, dusty working conditions. The House of Lords held that the pursuer could not, in his efforts to prove that the 'washing facilities' breach had caused his injury, rely on the 'but for' test, nor could he establish that the defender's omission had 'materially contributed' to his disease for it could not be ascertained *how* dermatitis was contracted. There was, nevertheless, medical evidence to the effect that cycling home covered in sweat materially increased the pursuer's *risk* of contracting dermatitis and, because the House of Lords equated this material increase in risk with a material contribution to the disease, causation was established. This risk-based approach was further developed in *Fairchild*.²⁶ As already mentioned, this was a case in which the court could not establish which employer's wrong caused the mesothelioma. Although the House of Lords knew that asbestos caused cancer, it did not know *how* this occurred and, because it could not rule out the 'single fibre theory', according to which the onset of mesothelioma could not be aggravated by further asbestos exposures, *Wardlaw* could not be invoked to establish cumulative causation on the part of both employers.²⁷ Because reliance on the 'but for' test would be unjust, the House of Lords decided to suspend the standard causation requirement, instead establishing causation on the basis that both employers had materially increased the risk of cancer. In reaching this conclusion, the court extended the *McGhee* principle. Whereas in that case the employer was held liable for damage he had definitely caused (via the hot working conditions and lack of showers) in *Fairchild*, the employers were held liable for damage that they *may not* have caused. In setting out this exception to the 'but for' test, the House of Lords prescribed various criteria for its application. For instance, it was said that the exception only applied if it was

²² T Honoré, 'Necessary and Sufficient Conditions in Tort Law' in DG Owen (n 14) 364.

²³ Wright, 'Causation in Tort Law' (n 2) 1782-1784.

²⁴ *Wardlaw* (n 1) 32.

²⁵ *McGhee* (n 10).

²⁶ *Fairchild* (n 10).

²⁷ *ibid* [7].

impossible to scientifically determine the cause of the plaintiff's injury and if the sources of harm were the same agent.²⁸

The *Fairchild* exception, though perhaps fair in the circumstances, is a rather flawed approach to causation, mostly because it was based on policy reasoning.²⁹ For instance, Lord Bingham argued that the injustice which would result from a strict application of the 'but for' test could justify a relaxation of the requirement to prove causation. In so doing, he was effectively abandoning the 'but for' test for normative reasons, in other words, because he felt the tortfeasors *should* be liable.³⁰

The overall object of tort law is to define cases in which the law may justly hold one party liable to compensate another. Are these such cases? A and B owed C a duty to protect C against a risk of a particular and very serious kind. They failed to perform that duty. As a result the risk eventuated and C suffered the very harm against which it was the duty of A and B to protect him. Had there been only one tortfeasor, C would have been entitled to recover, but because the duty owed to him was broken by two tortfeasors and not only one, he is held to be entitled to recover against neither, because of his inability to prove what is scientifically unprovable. If the mechanical application of generally accepted rules leads to such a result, there must be room to question the appropriateness of such an approach in such a case.

It is respectfully submitted here that when judges create exceptions to the 'but for' test on the sole basis of justice and fairness, they undermine the very purpose of causation: limiting liability. Lord Nicholls arguably missed the point of causation when he said that the 'but for' test should sometimes be relaxed when it became 'over-exclusionary.'³¹ On the contrary, it is this author's view that a causation test *should* be exclusionary. Failure to establish factual causation should not simply be some factor to be weighed against normative considerations; rather, it must guarantee the failure of the delictual claim. That is why the *Fairchild* exception, in allowing normative considerations to infiltrate the causation enquiry, falls foul of the 'limiting liability' criterion.

Closely related to this problem is the arbitrary nature of the rule. The facts of *Fairchild* were not special enough to merit exceptional treatment. Indeed, many cases have failed on account of an evidentiary difficulty;³² cases in which the 'injustice' reasoning used in *Fairchild* would have been equally valid.³³ Even the attempt in *Barker v Corus*³⁴ to provide a more principled justification for *Fairchild* did nothing to explain why the 'but for' test could be suspended in certain circumstances. Lord

²⁸ *ibid* [170].

²⁹ Although *Fairchild* was mostly driven by policy considerations, Lord Rodger did try to provide a principled justification for the decision, arguing that the principle in *McGhee* was directly applicable to the facts of this case. See generally his decision at [119]-[171], in particular [169]-[170], where he explains why the present case falls within the scope of the *McGhee* principle.

³⁰ *Fairchild* (n 10) [9].

³¹ *ibid* [40].

³² For example, *Wilsher v Essex Health Authority* [1988] AC 1074.

³³ J Morgan, 'Causation, Politics and Law: The English-and Scottish-Asbestos Saga' in Goldberg (n 4) 61.

³⁴ *Barker v Corus* [2006] 2 AC 572 (hereinafter '*Barker*').

Hoffmann argued that, in *Fairchild*, the relevant damage was not the mesothelioma but the *risk* of contracting it.³⁵ However, he also noted that the *Fairchild* exception only applies where the disease actually eventuated,³⁶ so it seems somewhat contradictory to characterise the risk as a form of damage in itself.

The largely policy-based, often flawed, reasoning employed in cases like *Fairchild* and *Barker* has major implications for the effectiveness of the causation enquiry. If there is no solid principle underlying an exceptional rule, its scope will be uncertain. Because it is a judicial precedent, the *Fairchild* rule will always be liable to extension by analogy,³⁷ yet arguably, owing to the dubious basis for creating the exception in the first place, its development will be haphazard and chaotic. If there is to be an exception to the 'but for' test, it must be completely certain when it applies, otherwise the causation enquiry will lack clarity and consistency.

In sum, the UK (and therefore Scottish) approach to factual causation falls short of all 'three criteria'. While the 'but for' test may lead to unfair results, the alternative tests compromise the very purpose of causation: limiting liability in a clear-cut way. The problem stems from having a harsh general rule and ill-defined exceptions, when what Scots law really needs is a single test of causation; one that will secure fairer results on a more principled basis. This article now proceeds to critically analyse the suitability of NESS as a replacement for the 'but for' test.

3. Is NESS a Rational Account of Causation?

According to Wright's NESS test, a condition X will be a cause of outcome Y if it was a necessary element in a set of actual antecedent conditions jointly sufficient for Y's occurrence.³⁸ Although Wright subsequently modified this definition, stating that X must be necessary *for the sufficiency* of a sufficient set,³⁹ the meaning is very much the same. For instance, a court examining whether a driver's negligence caused a pedestrian's injury would first construct a set of conditions sufficient for the injury to occur. These conditions would consist of facts that actually transpired, including the pedestrian's presence on the road and the driver's failure to pay attention. The court can establish causation if, absent the driver's negligence, this set of conditions would not be sufficient for the accident. It is significant that, in choosing the facts it wants to be part of the sufficient set, a court can '[rope] off' any facts it wishes to ignore.⁴⁰

Although the thrust of this part will be largely theoretical in character, focusing on whether NESS provides a rational account of causation, it will become clear in the next part that the test's ability to explain causation logically is crucial to its practical utility. Wright has claimed that his test captures the essence of

³⁵ *ibid* [35].

³⁶ *ibid* [48].

³⁷ Morgan (n 33) 59.

³⁸ Wright, 'Causation in Tort Law' (n 2) 1774.

³⁹ R Wright, 'Once More Into the Bramble Bush: Duty, Causal Contribution, and the Extent of Legal Responsibility' (2001) 54 *Vanderbilt Law Review* 1071, 1102-1103.

⁴⁰ R Wright, 'Acts and Omissions as Positive and Negative Causes' in JW Neyers, E Chamberlain and SGA Pitel (eds), *Emerging Issues in Tort law* (Hart Publishing 2007) 296.

causation,⁴¹ and it is this author's view that there may be some truth to this. Certainly, it seems promising that a condition need only be necessary to make a sufficient set of conditions. It is thanks to this concept of 'weak necessity' that X can be seen to cause Y, even though Y would still have occurred 'but for' X.⁴² However, Wright has possibly exaggerated NESS' theoretical strengths and so, in assessing NESS' ability to rationalise unusual forms of causation, this part aims to put his ambitious claims to the test.

A. 'Over-determination'

It is not a matter of controversy that NESS can detect causation in cases of over-determination, that is, cases in which multiple sufficient sets exist for the same outcome. Consider, for example, the scenario of two vehicles (V1 and V2) striking someone simultaneously.⁴³ Neither vehicle was necessary for the person's death, but if V1 is viewed as part of a set of conditions sufficient for the death, which does not include V2, V1 can be seen as a necessary element of that set. The same reasoning can be employed to hold that V2 was a necessary element of a set of conditions that does not include V1.⁴⁴ Therefore, the NESS test allows a court to identify both vehicles as causes. Similarly, if two separate fires burn down a house, both can be seen as necessary elements of their own sufficient sets.⁴⁵

A less straightforward example of over-determination is that of 26 separate discharges of chemicals into a river, which combine to kill cattle.⁴⁶ For the sake of example, it will be assumed that each discharge contains 1 unit of pollution and that a total of 18 units was needed to kill the cattle. This means that none of the discharges were independently necessary or sufficient for the damage caused. Nevertheless, all of these discharges played a causal role. Each one was, on a strict application of the test, a NESS.⁴⁷

In order to establish that a particular discharge was a NESS, a court would simply combine that discharge (1 unit) with 17 further discharges (1 unit each), thereby constructing a set that is *just sufficient* to kill the cattle.⁴⁸ In this particular sufficient set, the discharge of interest is a necessary element. Such an analysis has been criticised as artificial on the basis that it involves disaggregating the other 8 discharges from the set.⁴⁹ However, it is argued that this is as valid an application of the NESS test as the 'two vehicles' example above, a case that also involves constructing a sufficient set that artificially excludes the presence of the other condition. The 'pollution' example simply entails the same logic on a larger scale, there being far more sufficient sets to consider. As Miller explains, simply because

⁴¹ Wright, 'Causation in Tort Law' (n 2) 1805.

⁴² H Spector, 'The MMTS Analysis of Causation' in Goldberg (n 4) 339-340.

⁴³ This scenario is borrowed from Hogg. See Hogg, 'Developing Causal Doctrine' (n 8) 47-48.

⁴⁴ *ibid* 48.

⁴⁵ Miller, 'NESS for Beginners' (n 8) 324.

⁴⁶ This example is based on the case of *Warren v Pankhurst* 92 NYS 725 (NY Sup Ct 1904), *aff'd*, 93 NYS 1009 (AD 1905), *aff'd* 78 NE 579 (NY 1906) (as cited in *ibid* 327).

⁴⁷ *ibid* (Miller) 327.

⁴⁸ Wright, 'Causation in Tort Law' (n 2) 1793.

⁴⁹ Fischer (n 5) 290-292.

there are hundreds of sufficient sets, and the discharge in question is only a necessary element in some of them, does not preclude it from being a NESS.⁵⁰

If the pollution example is altered so that, instead of having 26 insufficient contributions, there is now one large discharge (20 units) and 16 smaller discharges (1 unit each),⁵¹ a new difficulty for the NESS test arises, known as ‘asymmetric’ over-determination.⁵² Since the number of units required to kill the cattle is 18, the large discharge (20 units) is independently sufficient. Establishing that one of the 16 smaller, insufficient contributions is a NESS means constructing a sufficient set that includes that contribution (1 unit) and supplementing it with *no more than* 17 additional units, some of which can only come from the larger discharge (20 units).⁵³ Because the larger discharge contains 20 units, it can only be used to complete a minimally sufficient set if the excess units can be disaggregated, yet to do so seems excessively artificial as it involves notionally dividing up a single entity (the large discharge) into, say, 17 units and 3 units respectively. Nevertheless, strained as this logic may appear, it is correct. NESS simply requires that a sufficient set consist of actual conditions and that is why the test *can* lead to a finding of causation in asymmetric over-determination cases.⁵⁴ Here, the large source *actually* contained 17 units of pollution (along with another 3) and so these 17 units can be combined with the smaller (1 unit) contribution to make the latter a NESS.

B. ‘Pre-emption’

Wright has claimed that, in certain cases involving potential causes X and Y, NESS supports the conclusion that X pre-empted Y, rendering it causally irrelevant.⁵⁵ One problematic example provided by Wright involves a ship whose access to a port has been blocked by two collapsed bridges. There are no alternative access routes. According to Wright, the first bridge to block the ship (B1) was a NESS of the ship’s delay but the bridge further upstream (B2) was not.⁵⁶ It would only have constituted a NESS if the ship had reached it. However, Fumerton and Kress think that the test,

⁵⁰ Miller, ‘NESS for Beginners’ (n 8) 327.

⁵¹ This is based on a similar modification of the pollution example made by Wright. See Wright, ‘Causation’ (n 2) 1793-1794.

⁵² MS Moore, *Causation and Responsibility: An Essay in Law, Morals and Metaphysics* (Oxford University Press 2009) 489.

⁵³ Wright, ‘Causation in Tort Law’ (n 2) 1793-1794.

⁵⁴ *ibid.*

⁵⁵ *ibid* 1795. Here, Wright gives an example of ‘pre-emption’ in which P drinks tea which has been poisoned by C but, before the poison takes effect, P is shot dead by D. According to Wright, D’s gunshot pre-empts C’s poisoning of the tea, which is to say that D’s gunshot was a NESS of P’s death and C’s poisoning of the tea was not. This is because, says Wright, in order to create a sufficient set of actual conditions which included, as a necessary element, C’s poisoning, P must have been alive when the poison took effect. The essence of Wright’s concept of ‘pre-emption’ is that, even if there is an *actual* condition (in this case P’s consumption of the poisoned tea) which, together with other actual conditions (e.g. the absence of an antidote), is sufficient to *guarantee* the occurrence of a particular outcome (namely P’s death), that condition is nevertheless not a cause because of some other condition (in this case D’s gunshot). While ‘over-determination’ means that *both* conditions X and Y are causes, ‘pre-emption’ means that X prevents Y from having any causal effect.

⁵⁶ *ibid* 1796-1797.

properly applied, renders both bridges NESSs.⁵⁷ In order to appreciate their criticism, it is necessary to examine the *type* of sufficiency to which NESS relates.

If event X is *lawfully* sufficient for event Y it means that, if X is present, Y's presence will be guaranteed by some law of nature.⁵⁸ For instance, fire is lawfully sufficient for oxygen because the presence of fire guarantees the presence of oxygen. Fumerton and Kress are of the view that NESS embodies this notion of lawful sufficiency; that a 'NESS' cause is a necessary element of a *lawfully* sufficient set.⁵⁹ The problem with this is that lawful sufficiency does not necessarily denote a causal relationship. For instance, a barometer falling guarantees the presence of a storm but it does not follow from this that the barometer *causes* the storm.⁶⁰ Nevertheless, lawful sufficiency is all NESS requires and, because the collapse of B2 guarantees that the ship will not reach the harbour, B2 is part of a lawfully sufficient set.⁶¹ Contrary to what Wright argues, the sufficiency of this set does not depend on the ship arriving at B2.

The only means of holding that B2 was not a NESS is by invoking a concept of causal sufficiency.⁶² Causal sufficiency is a special type of lawful sufficiency, according to which the presence of X will be sufficient to *cause* Y. Thus, in the same way that a barometer falling is not causally sufficient for a storm, Wright might be able to argue that B2, though lawfully sufficient for the ship's delay, would only be causally sufficient if the ship reached B2, which did not occur. However, acceptance of this approach would render NESS viciously circular. In defining causation, Wright would be relying on a concept of sufficiency itself predicated on a certain understanding of causation, which, to paraphrase Fumerton and Kress, would effectively amount to defining causation as 'causation.'⁶³ Wright's subsequent attempts to justify 'causal sufficiency' have been unconvincing.⁶⁴ For example, he has said that, although B2 guarantees that the ship will not arrive at the harbour, it does not guarantee that the ship's delay will be 'caused by' B2.⁶⁵ The use of the phrase 'caused by' only serves to further illustrate the circularity of his theory. Wright is effectively saying that B2 was not a *cause* of the ship's delay because it did not guarantee that the ship's delay would be *caused* by B2. In a recent article, Wright insisted that causal sufficiency is not circular by clarifying that it can be established by means of empirical tests:⁶⁶

Our knowledge of the required conditions in the antecedent of a causal law-and thus of the direction of causation-is based on experience and empirical observation, by ourselves or others. Scientists employ Mill's Difference Method in carefully designed experiments to see if the non-instantiation of a supposed antecedent condition makes a difference in the occurrence of the

⁵⁷ Fumerton and Kress (n 6) 100-101.

⁵⁸ *ibid* 92-93.

⁵⁹ *ibid* 94.

⁶⁰ *ibid* 93.

⁶¹ *ibid* 100-101.

⁶² *ibid* 101-102.

⁶³ *ibid* 102.

⁶⁴ Stapleton (n 3) 477.

⁶⁵ Wright, 'Acts and Omissions' (n 40) 305.

⁶⁶ Wright, 'The NESS Account of Natural Causation' (n 7) 289.

consequence. For example, we determine by observation or experimentation that eliminating a flagpole or changing its height eliminates or changes the length of the flagpole's shadow, but not vice versa.

However, as Steel has pointed out, this would not render NESS any less circular.⁶⁷ In applying a scientific test to determine what is causally sufficient, one would inevitably have certain preconceptions of what is causally sufficient.

Fumerton and Kress' reasoning applies to all 'pre-emption' cases, including those concerning omissions. A key example involves the renter of a car neglecting to repair the brakes and an accident occurring when the driver, not knowing they are faulty, fails to apply them anyway.⁶⁸ Wright has argued that the driver's omission to apply the brakes pre-empts that of the renter to repair them.⁶⁹ However, contrary to Wright's conclusion, both omissions are NESSs. The failure to repair was part of a lawfully sufficient set because it guaranteed that the car would crash. Wright's argument, namely that the failure to provide a safeguard could only be an *actual* condition if an attempt was made to use it, is unconvincing, for the failure to repair the brakes *actually* occurred. Moreover, Wright's reasoning fails to demonstrate why it was the driver's omission that pre-empted that of the renter and not vice versa. As Fischer points out, the same logic can be used to conclude that the driver's omission would only have a causal effect if the brakes were working.⁷⁰ In short, Wright provides no compelling reasons that explain why both omissions, that of the driver *and* that of the renter, are not NESSs. His assertion that the two factors are distinguishable because the driver's omission had causal priority,⁷¹ is based on the same circular notion of causal law that he invokes in the 'two bridges' example.⁷²

In sum, Fumerton and Kress have exposed a significant flaw in the NESS test. Until Wright can provide a non-circular explanation for why the driver's omission pre-empts that of the renter and why B2 is any less a NESS than B1, these scenarios, and indeed all purported cases of pre-emption, will be indistinguishable from cases of over-determination.⁷³

C. 'Indeterminate Causation'

A particularly difficult issue is NESS' applicability to cases involving scientific indeterminacy.⁷⁴ The problem is that, where there are multiple exposures to a dangerous chemical such as asbestos and it is not known *how* exactly this causes mesothelioma, it is impossible to construct a sufficient set of actual conditions that will include, as a necessary element, only *one* employer's wrongdoing. For instance, if an individual has been exposed to asbestos by three consecutive employers (E1, E2

⁶⁷ S Steel, 'Publication Review: Perspectives on Causation' (2012) 16(3) ELR 458, 460.

⁶⁸ This example is based on *Saunders System Birmingham Co. v Adams* 217 Ala 621 So 72 (1928) (as cited in Wright, 'Causation in Tort Law' (n 2) 1801).

⁶⁹ *ibid* (Wright) 1801.

⁷⁰ D Fischer, 'Causation in Fact in Omission Cases,' (1992) Utah Law Review 1335, 1358-1360.

⁷¹ Wright, 'Once More' (n 39) 1130-1131.

⁷² Fischer, 'Insufficient Causes' (n 5) 312.

⁷³ For example, the 'two fires' scenario: see text accompanying footnote 45.

⁷⁴ For an example of scientific indeterminacy, see text accompanying footnote 21.

and E3), there are various possible combinations of exposures that could have precipitated the disease. It is possible that E1's exposure caused the disease and those of E2 and E3 played no role in aggravating it or, alternatively, that all three exposures played a role. However, there is no way of verifying which possibility transpired. It is, therefore, impossible to construct a set that includes, say, E1's exposure but not those of E2 and E3. There is no way of guaranteeing that that set was sufficient. It is for that reason that NESS cannot lead to a finding of causation in cases of scientific uncertainty.⁷⁵

In conclusion, it cannot be denied that NESS, in its recognition that there can be multiple combinations of conditions sufficient for one outcome,⁷⁶ can detect a far wider category of causation than the 'but for' test. However, NESS suffers from several theoretical drawbacks. Particularly devastating is Wright's insistence on a concept of causal, rather than lawful, sufficiency, which has led to accusations that the test is viciously circular.⁷⁷ This, coupled with NESS' related inability to explain pre-emption and indeterminacy cases, suggests that it is far from being the philosophically comprehensive account of causation that Wright claims it to be. However, even if the test is imperfect, it may still be useful on a more practical level.

4. Would NESS be of Practical Use to the Scottish Courts?

The theoretical flaws identified in part III of this article might be fatal to Wright's claim that NESS captures the essence of causation,⁷⁸ but they are by no means fatal to the test's utility. As Stapleton explains, the law does not need a philosophically sound account of causation but can rather 'choose' a certain meaning to suit its needs.⁷⁹

(...) causal language can be used to express information from a variety of interrogations into our world pursued for different purposes: and it is only once we have chosen which is the underlying interrogation in our dialogue that we can infuse our causal language with unambiguous meaning. Thus, for example, it is only because scientific method requires scientists to expose their choice of interrogation by explicitly recording the parameters of their enquiry, that scientific discourse can proceed unambiguously (...) [W]hilesoever philosophers ignore the necessity for them to agree on an interrogation ((...) be it prevention or blame or explanation etc) to underlie their use of causal language, their discussions of "the concept of causation" will be doomed to proceed at cross-purposes.

Because the term 'cause' can be used to convey different types of information, ranging from mere physical involvement to the attribution of blame, the law can, and

⁷⁵ Hogg, 'Developing Causal Doctrine' (n 8) 48.

⁷⁶ C Miller, 'Causation in personal injury after (and before) *Sienkiewicz*,' (2012) 23(3) Legal Studies 396, 399.

⁷⁷ Stapleton (n 3) 472.

⁷⁸ Wright, 'Causation in Tort Law' (n 2) 1789.

⁷⁹ Stapleton (n 3) 439 (footnote 15 of that article).

should, specify a particular interrogation to underlie the causal enquiry. The upshot of this is that NESS can still be useful to the law, even if it does not 'engage philosophers'⁸⁰ or comply with some 'divine' standard.⁸¹ That is not to guarantee, however, that NESS *would* be of use to the courts but simply that the law can 'choose' what it means by causation.⁸² The following sections will investigate, with close reference to the 'three criteria' identified in part II, whether NESS would make a good 'choice' for Scotland.

A. Practical Benefits

The primary advantage of adopting NESS would be fairness in the law as, unlike the 'but for' test, NESS can rationalise cases of over-determination. Stapleton has argued that, in order for the causal enquiry to meet the law's needs, factor X should be deemed a cause of outcome Y if it was 'involved' therein⁸³ and further, that NESS, all theoretical imperfections notwithstanding, would be an ideal algorithm for assessing 'involvement'.⁸⁴ Stapleton's concept of 'involvement' takes three forms, namely necessity,⁸⁵ duplicate necessity⁸⁶ and contribution,⁸⁷ yet the 'but for' test can only detect the first type. Like the 'but for' test, NESS can recognise involvement by necessity but, as part III of this article has demonstrated, it also identifies causation in cases of duplicate necessity⁸⁸ and contribution.⁸⁹ It is because the 'but for' test does not recognise 'involvement' in all its forms that it begets such harsh decisions. NESS, with its ability to identify a broad range of causal 'involvement', would secure far fairer results for pursuers.

For instance, NESS may have allowed a finding of causation in the somewhat harsh decision of *Wilsher v Essex HA*.⁹⁰ Although judges often compare this case to *Fairchild*, the fact that an excess of oxygen was only one of several factors that could

⁸⁰ Miller, 'NESS for Beginners' (n 8) 337.

⁸¹ M Hogg, 'The Role of Causation in Delict' (2005) *Juridical Review* 89, 113.

⁸² Stapleton (n 3) 438-441.

⁸³ *ibid* 443-446.

⁸⁴ *ibid* 474.

⁸⁵ *ibid* 441. For the avoidance of confusion, Stapleton's concept of 'necessity' simply describes a typical 'but for' relationship e.g. X's gunshot was necessary for Y's injury because, if X had not shot Y, Y would not have been injured.

⁸⁶ *ibid* 441-443. Stapleton's concept of 'duplicate necessity' is a particular form of over-determination. The aforementioned 'two hunters' scenario (see text accompanying footnote 20), 'two vehicles' scenario (see text accompanying footnotes 43 & 44) and 'two fires' scenario (see text accompanying footnote 45) are all examples of duplicate necessity because if, say, in the 'two vehicles' scenario, V1 is notionally removed, V2 now becomes necessary for the pedestrian's injury and vice versa.

⁸⁷ *ibid* 443-444. Stapleton's concept of 'contribution' is simply an example of over-determination. Both of the aforementioned pollution examples (see text accompanying footnotes 46 and 51) are instances of 'contribution'. 'Contribution' is a wider form of over-determination than 'duplicate necessity'. In the 'two vehicles' example, V1 is independently sufficient (along with all other relevant factors e.g. the pedestrian's presence on the road) for the pedestrian's injury. However, in the pollution examples, a particular discharge of 1 unit will need to be notionally combined with some of the other discharges to make a sufficient set.

⁸⁸ For example, the 'two fires' scenario. See text accompanying footnote 45.

⁸⁹ For example, the 'pollution' scenarios. See text accompanying footnotes 46 and 51.

⁹⁰ *Wilsher* (n 32).

have caused retrolental fibroplasia (RF), suggests that this was a potential case of over-determination. Had NESS applied, the court may have been able to construct a sufficient set for RF including, as a necessary element, the excess of oxygen, and excluding all competing factors. Lord Bridge was somewhat equivocal as to whether an excess of oxygen was sufficient for RF⁹¹ but if, on closer examination of the scientific evidence, he had found that an excess of oxygen *was* part of a set of conditions sufficient for RF, causation would have been established and a fairer result would have been obtained.

What makes NESS particularly attractive, however, is its ability to achieve justice on a logical basis, without recourse to normative considerations, legal fictions or policy.⁹² Because judges would be required to confine their analysis of causation to empirical concepts, namely necessity and sufficiency, determinations of whether or not particular conditions were NESSs would be relatively unambiguous. As a result, the law would be consistent. There would be no need to depart from the NESS test or introduce ill-defined exceptions to supplement and attenuate the general rule, for the general rule would be perfectly satisfactory. Thus, whether examining the simplest 'but for' cases or the most complex cases of over-determination, the courts could rely on a single test of factual causation.

Indeed, if NESS had been the standard test at the time of *Wardlaw*,⁹³ the presence of multiple concurrent sources would not have been an issue.⁹⁴ The court would simply have combined the 'negligent' dust with *just enough* of the 'non-negligent' dust to construct a sufficient set. The 'negligent' dust would have been deemed a NESS of the pursuer's injury and so there would have been no need to devise an exception based on 'material contribution'. Indeed, it would be preferable to analyse cases like *Wardlaw* in terms of NESS. In contrast to the vague concept of 'material contribution', the meaning of which is apt to vary according to judicial intuition, it would not be so easy for judges to differ on what constitutes necessity and sufficiency. Whereas 'material contribution' is a question of degree, the concepts of necessity and sufficiency are more objective and absolute. In short, because NESS can establish fair results on a more rational basis in over-determination cases,⁹⁵ it is suggested here that it satisfies the criterion of 'clarity and consistency'.

It almost goes without saying that NESS would be useful in over-determination cases. More problematic is its applicability to cases of indeterminate causation. It was established in part III that NESS does not provide a solution in such cases. Nevertheless, the test could recognise causation in such cases if used in a certain way. Stapleton has argued that NESS should be calibrated to give the most desirable results: 'I argue that NESS is merely an algorithm; an algorithm the catchment of which we design so that it will identify all forms of involvement of interest to the Law.'⁹⁶ Thus, although there is no rule inherent in the concept of a 'NESS' that would allow causation to be identified in cases like *Fairchild*,⁹⁷ it may

⁹¹ *ibid* 1091.

⁹² Miller, 'NESS for Beginners' (n 8) 323.

⁹³ *Wardlaw* (n 1).

⁹⁴ Miller, 'NESS for Beginners' (n 8) 327.

⁹⁵ Adams (n 20). See text accompanying footnote 19 of that article.

⁹⁶ Stapleton (n 3) 477.

⁹⁷ *Fairchild* (n 10).

nevertheless be possible to build such a rule into the test. This may be objected to on the grounds that it would allow normative considerations to influence *how* NESS is used. However, so long as the concepts *within* the test are empirical (i.e. necessity and sufficiency), it will still be possible to apply NESS in a clear-cut, objective manner.

It is submitted that, if NESS is adopted, the following rule could be built into the test: in establishing that the defender's wrongdoing was a NESS, the pursuer would first have to construct a set of conditions that was *definitely* sufficient, even if it was not minimally sufficient. Then, the court would notionally remove the defender's wrongdoing from the set and, if it was *no longer definitely* sufficient, the wrongdoing would be deemed a necessary element thereof. This would mean that, in the example above,⁹⁸ the employee could establish that each employer's exposure was a NESS of his or her injury. He or she would begin by constructing a sufficient set containing all three exposures. Assuming that the individual was only exposed to asbestos in the workplace, they could construct the following sufficient set: all three consecutive exposures taken together (i.e. those of E1, E2 and E3), along with any other relevant conditions such as a lack of safety precautions. This set is *definitely* sufficient. Now, in order to demonstrate that, say, E1's exposure was a necessary element of this set, the employee would simply have to show that, without this exposure, the set was *no longer definitely* sufficient. It is uncertain whether E1's exposure actually played any role in the contraction of mesothelioma but, equally, it cannot be ruled out that it *did*. In this way, removing E1 from the set does not necessarily make it insufficient; it simply casts the set's sufficiency into doubt. Yet, because of the special rule, this would suffice to establish that the exposure was a NESS. In short, a sufficient set containing conditions E1, E2 and E3, is the closest approximation of a definitely sufficient set that science will currently permit and so, pending some scientific discovery to the contrary, all three conditions will be deemed NESSs.

Various objections may be levelled against this rule. One might argue, for example, that the above is no different from using the 'but for' test and reversing the burden of proof, requiring the defender to show that his or her particular exposure was *not* necessary. However, this is not technically true. The pursuer would still have the burden of proof of causation, they would simply have to prove slightly less, namely that the removal of the condition casts the set's sufficiency into doubt. Moreover, with the 'but for' test, a 'reversal of the burden of proof' rule would exist as an exceptional one, only applying in cases of scientific indeterminacy. With the NESS test, on the other hand, the 'no longer definitely sufficient' rule, would simply be part of the test: it would apply in all cases but would only make a difference in cases of scientific indeterminacy.

Consider, for example, the 'two vehicles' scenario.⁹⁹ Here, the court would first construct a set including V1, V2 and the pedestrian's presence on the road. This is *definitely* sufficient for the accident. Without V1, the set is still *definitely* sufficient. However, the removal of V2 as well means that the set is *no longer definitely sufficient* and so V2 is a NESS. Of course, the court would also know that this set was *definitely*

⁹⁸ See text accompanying footnotes 74 & 75.

⁹⁹ See text accompanying footnotes 43 & 44.

insufficient without V1 and V2, but there would be no need for the pursuer to prove this. The salient point is that this result (both V1 and V2 are NESSs) is exactly the same result that would be obtained if the pursuer had to prove that the set was *definitely insufficient* without the two vehicles. The 'no longer definitely sufficient' rule therefore has no impact. The fact that this rule would apply to all cases, and not as an exception in cases involving scientific indeterminacy, would minimise confusion. This, combined with the fact that the rule is articulated in terms of an empirical concept, namely sufficiency, would ensure clarity in the law. Whether a particular set was 'not definitely sufficient' would always be a clear-cut matter.

If NESS is employed in the above way, the test would be able to provide fair results in cases involving scientific indeterminacy but on a more rational basis. For a start, there would be a more principled basis for joint and several liability.¹⁰⁰ In *Barker*,¹⁰¹ the House of Lords tried to justify mere several liability by characterising the *damnum* as the risk of contracting mesothelioma and, although the Compensation Act 2006¹⁰² has reversed the harm done by *Barker*, at least with regards to asbestos cases, this problem would never have arisen under the NESS test. NESS, used in the way proposed above, would have irrefutably established that the employers *caused* the *disease*, leaving it in no doubt that they were jointly and severally liable.

In sum, NESS' practical utility resides in its ability to establish fairer results on a more principled basis. It is a useful explanatory device capable of establishing causation in over-determination cases and, although not inherently capable of rationalising indeterminate causation, could do so in practice if supplemented by a special rule. There is no doubt that NESS satisfies the two criteria of fairness and 'clarity and consistency' but there is a flipside to this, namely the risk that, in adopting NESS, the Scots courts would make causation too easy to establish. The next section will focus, *inter alia*, on this potential problem and whether it outweighs the potential benefits of adopting NESS.

B. Potential Dangers

Some objections to the NESS test are so trivial that they can be dismissed almost immediately. For instance, Fumerton and Kress have argued that NESS, if it relies on a concept of lawful sufficiency, presupposes a deterministic universe and so could not address cases involving probabilistic sources of harm.¹⁰³ They provide the example of an 'indeterministic' bomb, which, depending on a random process, may explode in a minute or 2000 years.¹⁰⁴ The point is that, if the bomb detonates, injuring someone, it could not be established that the bomb was lawfully sufficient as it was not *guaranteed* to explode at that particular time. However, in practice, a judge would hold that the bomb was lawfully sufficient for the person's death because the bomb *did* go off, thus guaranteeing the person's injury on *this particular occasion*. Thus,

¹⁰⁰ Miller, 'NESS for Beginners' (n 8) 337.

¹⁰¹ *Barker* (n 34).

¹⁰² s 3.

¹⁰³ Fumerton and Kress (n 6) 97-98.

¹⁰⁴ *ibid*.

though NESS' presupposition of determinism may be a legitimate theoretical grievance, it presents no practical problems.

Similarly pedantic is Fumerton and Kress' argument that any fact can be considered a NESS.¹⁰⁵ They give the example of X starting a fire and argue that the fact that he was wearing a blue shirt, or any other irrelevant condition, can be rendered a NESS thereof. Although their logic, which consists of manipulating such truth functional operators as 'NOT' and 'OR', is perfectly sound, it is predicated on the existence of the following state of affairs: X was not wearing a blue shirt OR N (a state of affairs lawfully sufficient for the fire). By making the proposition 'X was not wearing blue shirt' and 'N' alternatives to one another, Fumerton and Kress are able to conclude that X's blue shirt was a NESS. However, this state of affairs of the structure 'Either...Or...' simply does not arise in practice: judges invariably define a set of facts according to what 'is' the case. As such, it is suggested here that Fumerton and Kress' criticism is purely academic.

On the other hand, the fact that NESS would make causation easier to establish does raise serious practical issues. For a start, in order for NESS to provide a non-circular account of causation, it would have to recognise as a cause any necessary element of a set that is *lawfully* sufficient. In other words, the set will be deemed sufficient if it *guarantees* the outcome's occurrence. This approach could be used to justify all manner of absurd results, for example, the determination that a barometer falling causes a storm.¹⁰⁶ Such results would make a farce of the causation requirement. Consequently, if NESS is to satisfy the 'limiting liability' criterion, it must recognise that the existence of lawful sufficiency between two events does not always denote a causal relationship. Of course, to draw such a conclusion would be to invoke a concept of causal sufficiency but the circularity that this would entail, although theoretically unsatisfying, would not necessarily be a problem in practice.

Then again, even if there is a way of ruling out the barometer as non-causal, the following question arises: can B2 (from the 'two bridges' example)¹⁰⁷ be deemed non-causal by the same token? These issues are problematic. On the one hand, a causation test must be able to limit liability in a significant way; on the other, a causation test must be clear-cut and there is a risk that, by relying on an undefined concept of causal sufficiency to hold that a barometer did not cause a storm and that a certain bridge did not cause a ship's delay, judges could never achieve consensus as to what constitutes a NESS. As mentioned earlier,¹⁰⁸ Wright's notion of causal sufficiency can be used to reach two completely different conclusions in the 'faulty brakes' problem, which raises the question: if causal sufficiency is such an ambiguous concept, could judges ever invoke it on a consistent basis?

Before examining the issue of causal sufficiency, however, it is worth observing one way in which NESS can hold that factors are causally irrelevant without invoking a concept of causal sufficiency, that is, by articulating the causal question more specifically.¹⁰⁹ Where the causal question is, 'what caused the ship's

¹⁰⁵ *ibid* 94-95.

¹⁰⁶ See text accompanying footnote 60.

¹⁰⁷ See text accompanying footnote 56.

¹⁰⁸ Fischer, 'Causation in Fact' (n 70) 1358-1359.

¹⁰⁹ Adams (n 20). See text accompanying footnote 27 of that article.

delay?’ both B1 and B2 are NESSs for, as discussed above, both are part of lawfully sufficient sets for that outcome. However, if the question is phrased as ‘what caused the ship to stop *behind* B1?’ then B2 becomes causally irrelevant as the fact that it had collapsed did not guarantee that the ship would stop *at that particular point*. Indeed, this technique of articulating the outcome in greater detail is not peculiar to the NESS test. As Adams makes clear, the same result can be achieved under the ‘but for’ test.¹¹⁰ But for the collapse of B1, the ship would not have stopped *just behind* B1, however this result would have occurred regardless of the collapse of B2.

Nevertheless, it is argued here that this sort of pre-emption is of limited utility. It hinges on the rather contrived device of specifying the outcome in minute detail, in effect altering the causal question to obtain the desired answer. This can only be done to a limited extent in legal practice. As Wright mentions:¹¹¹

The courts (...) do not qualify the consequence by specifying its non-salient details or the time, location or manner of its occurrence when describing or applying the *sine qua non* analysis.

Wright’s point, although perhaps made with the American courts in mind, is certainly true of the Scots courts, which are far more likely to articulate the causal question as: ‘would the pursuer *have been delayed* had B1 not collapsed?’ than ‘would the pursuer have *had to stop x metres from* B1 had B1 not collapsed?’ More significantly, phrasing the causal question more narrowly does not provide a solution to the ‘barometer’ problem.

A more practical solution would be to have a rebuttable presumption that, if X was a necessary part of a *lawfully* sufficient set for Y, X caused Y. However, if the court can be satisfied, on the balance of probabilities, that the relationship between X and Y could not possibly have been causal then the court would be allowed to impose a requirement of causal sufficiency and hold that X was *not* a cause of Y on that basis. Thus, although there would be a presumption, under the NESS test, that the fall of a barometer caused a storm, this presumption could be rebutted because it is well known that the fall of a barometer cannot cause storms and that both of these events are the *effects* of a drop in air pressure. The existence of a presumption would ensure certainty in the law. This presumption could only be overturned in extreme cases such as that of the barometer, the sort of case in which there is clear-cut, unequivocal evidence that a particular relationship is not causal. In the bridges example, it would not be possible to rebut the presumption that lawful sufficiency was enough for the purposes of causation as there is not really any overwhelming evidence to suggest that the relationship between B2 and the ship’s delay was not causal.

To the purist and the philosopher, any recourse to a notion of causal sufficiency is circular, illogical and out of the question. From the standpoint of a judge, however, such an analysis would be acceptable. This is because all judges, even when deciding cases in terms of the ‘but for’ test, presuppose a certain notion of

¹¹⁰ *ibid.* See text accompanying footnote 24 of that article.

¹¹¹ Wright, ‘The NESS Account of Natural Causation’ (n 7) 294.

causal sufficiency. For example, the court in *Barnett v Chelsea*,¹¹² in assessing whether a doctor's negligence was a cause of an individual's death, were effectively asking the following question: 'in a hypothetical world in which the doctor had *not* refused to treat the individual, would all the remaining actual conditions, most notably the arsenic poisoning, have been *sufficient* to *cause* his death?' This notion of causal sufficiency, tacitly invoked, was not one on which the judges were likely to differ, it being a well-established scientific fact that the relationship between arsenic poisoning and death is not simply lawful but causal. Conversely, judges would likely all agree that the symptoms of a disease, although lawfully sufficient for the disease in that they guarantee its presence, were not sufficient to cause the disease. They would know, based on scientific evidence, that it was the disease that caused the symptoms, not vice versa.

In short, if NESS is used in a particular way it could adequately rule out certain conditions as causally irrelevant in extreme cases, allowing it to satisfy the 'limiting liability' criterion. However, this must not come at the expense of clarity in the law. As Thomson has observed, if words like 'sufficiency' are not clearly defined, this can lead to confusion.¹¹³ With this point in mind, it must be clear to judges in what exact circumstances NESS requires mere lawful sufficiency and when NESS requires causal sufficiency. There is thus a need for a presumption that lawful sufficiency is enough; a presumption that can only be overturned if the lack of causal sufficiency is so overwhelmingly evident that no judge is bound to disagree about it. This goes to show that even if NESS cannot, in strict theoretical terms, employ a concept of causal sufficiency to brand certain conditions causally irrelevant, it can do so, to a limited extent, in practice.

Aside from the above problem of lawful sufficiency, some commentators feel that the NESS test cheapens the notion of causation in a more fundamental way.¹¹⁴ It has been argued, for example, that the focus on sufficiency rather than necessity makes it too easy to establish causation in cases involving multiple wrongdoers. Chris Miller has commented that the NESS test's utility lies in its recognition that there can be various causal pathways 'capable' of triggering a certain outcome, but this may also be its fundamental flaw.¹¹⁵ The word 'capable' is significant. It suggests that NESS is concerned not so much with what did happen as with what *could* have happened. Just as the 'but for' test, with its emphasis on necessity, is overly restrictive, undue emphasis on sufficiency could lead to over-inclusiveness.¹¹⁶

This problem is particularly acute in cases of 'contribution'.¹¹⁷ In a way, these cases represent the controversial outer limits of NESS causation, where a factor can play a seemingly minor role in a particular outcome but still be deemed a cause. If anything is going to render the causal requirement meaningless or impotent, it is these cases. Fischer gives the extreme example of X adding a teaspoon of water to a

¹¹² *Barnett v Chelsea and Kensington Hospital Management Committee* [1969] 1 QB 428.

¹¹³ JJ Thomson, 'Some Reflections on Hart and Honoré, *Causation in the Law*' in MH Kramer, C Grant, B Colburn and A Hatzistavrou (eds) *The Legacy of HLA Hart: Legal, Political and Moral Philosophy* (Oxford University Press 2008) 143.

¹¹⁴ Fischer, 'Insufficient Causes' (n 5) 290-291.

¹¹⁵ Miller, 'NESS for Beginners' (n 8) 327.

¹¹⁶ E Quill, 'Successive Causes and the Measurement of Damages' (2002) 37 Irish Jurist 91, 97.

¹¹⁷ For a definition of Stapleton's concept of 'contribution' see footnote 87.

tsunami.¹¹⁸ The NESS test makes this tiny contribution of water causally relevant by notionally combining it with just enough tsunami water to make a sufficient set for a particular piece of damage. Applying this analysis to each piece of damage, the teaspoon of water can be deemed a NESS for all the destruction wrought by the tsunami. Because of this absurd result, Fischer feels that the NESS test should not make a finding of causation in such cases.

To an extent, Fischer is correct. It would be nonsensical to hold X liable for any of the tsunami damage. However, even if X were held a cause, the various normative filters, most notably breach of duty and legal causation, would ensure that X was not held liable for such a trivial contribution. Moreover, there would be very good reasons for allowing the teaspoon of water to be held a cause. To modify Fischer's example,¹¹⁹ imagine that thousands of individuals each simultaneously dispense a teaspoon of water into a river, flooding someone's home. In the interests of fairness, the law requires a concept of causation flexible enough to recognise each person's contribution to the flood as a cause, otherwise everyone would be able to hide behind the others' wrongdoing. At the same time, if NESS is embraced, it must be embraced wholeheartedly, which is to say that the empirical concept of a NESS must never yield to normative considerations. There is a need to maintain clarity and consistency in the law and, to that end, if NESS detects causation in less extreme cases of contribution it should also detect causation in anomalous cases such as the tsunami example. Fischer insists that it should be possible to hold that a certain factor, though a NESS, was nevertheless not a cause because it was too trivial.¹²⁰ However, such an approach would adulterate the empirical purity of the NESS test and result in a lack of clarity in the law. It is suggested that the price of consistency and fairness may be the occasional absurd finding that a teaspoon has caused a tsunami but, so long as this sort of case can be handled by normative filters like legal causation, no practical harm should result.

On a more general level, it is worth pointing out that NESS does not render causation completely meaningless. The concepts of necessity and sufficiency may sometimes lead to absurd results, but they have the advantage of being empirical concepts that leave no room for normative issues or judicial discretion. In addition, there are plenty of circumstances in which NESS would *not* lead to a finding of causation. In *Kay's Tutor*,¹²¹ for example, no sufficient set for deafness could have been constructed that included, as a necessary element, a penicillin overdose, as there was no evidence that such an overdose could cause deafness. Thus, the defenders' negligence would not have been deemed a NESS. The NESS test may make causation easier to establish, but it still meets the primary 'criterion' of a useful test: limiting liability in a meaningful way. If adopted, the test would present something of a challenge to the causal minimalists' argument that the causation enquiry is an

¹¹⁸ Fischer, 'Insufficient Causes' (n 5) 290-291.

¹¹⁹ This modification is loosely based on a similar modification made by Wright. See Wright, 'The NESS Account of Natural Causation' (n 7) 304-305.

¹²⁰ Fischer, 'Insufficient Causes' (n 5) 288.

¹²¹ *Kay's Tutor* (n 18) 167.

illusory test that is easily satisfied.¹²² Far from it, the NESS test would present an uncompromising obstacle to many claims.

Some objections raised against NESS concern its lack of practicability and one key problem in this regard is that of the 'proliferation of NESSs'.¹²³ Moore has argued that the problem with NESS, and sufficiency theories in general, is the 'sheer number of events and states of affairs' needed to create a sufficient set.¹²⁴ He bemoans the fact that, in order for him to compile a list of conditions jointly sufficient for him to write his book, he would need to mention the neurological processes in his brain, his room being sufficiently warm so that his fingers did not freeze, Caesar crossing the Rubicon, the Big Bang and an almost infinite amount of other NESSs.¹²⁵

In addition, a truly sufficient set consists of an infinite amount of negative NESSs, that is, any event the non-occurrence of which was necessary for the set's sufficiency. Thus, in the 'two hunters' example,¹²⁶ the fact that the victim was not wearing a bullet-proof vest, that the guns were not loaded with marshmallows rather than bullets, that Bambi's mother did not jump in the line of fire just in time, are all NESSs of the victim's death. In short, for a pursuer to establish, to the point of certainty, the existence of a sufficient set, he or she would have to list all the NESSs of which it was composed, a nigh impossible task. However, in legal practice, the pursuer would only have to establish the existence of a particular sufficient set on the balance of probabilities. More significantly, the large number of positive NESSs (e.g. the existence of gunpowder, the invention of the gun years ago) can usually be expressed by much smaller groups of NESSs. Thus, the NESSs concerning the minutiae of the gun (e.g. the fact that it was loaded, working etc.) and how it was shot (e.g. the distance from which it was shot, the trajectory of the bullet, the fact that the bullet hit the victim etc.) can be neatly packaged into a single NESS: 'hunter 1 shot the victim'. As for the omission NESSs, the very fact that the result in question, namely the victim's death, has occurred suggests that all the 'omission' NESSs were operating at the relevant time. Similarly, both Wright and JS Mill favour the strategy of summarising the vast multitude of omission NESSs as 'the absence of any preventing cause'.¹²⁷ There are, therefore, means of applying the test in a practical, efficient way.

Another potential problem regarding practicability is NESS' supposed complexity, on account of which judges have been reluctant to embrace the test.¹²⁸ However, as Hogg points out, the test is not inherently complex and its daunting nature can probably be attributed to the 'fanciful examples' favoured by academics to illustrate its application.¹²⁹ Thus, the test is only as complex and mysterious as the academics choose to make it and, true to his point, Hogg uses simple,

¹²² R Wright, 'The Nightmare and the Noble Dream: The Road not taken' in Kramer, Grant, Colburn & Hatzistavrou (n 113) 177.

¹²³ Fumerton and Kress (n 6) 98-99.

¹²⁴ Moore (n 52) 477.

¹²⁵ *ibid.*

¹²⁶ See text accompanying footnote 20.

¹²⁷ Wright, 'Acts and Omissions' (n 40) 291.

¹²⁸ Hogg, 'Developing Causal Doctrine' (n 8) 47-48.

¹²⁹ *ibid.*

comprehensible examples, to illustrate how the test works.¹³⁰ Moreover, given the wide range of causation that NESS can detect, it arguably makes causation far simpler than it would be if, say, a judge had to apply three separate tests of 'necessity', 'duplicate necessity' and 'contribution'.¹³¹ There is, however, one way in which NESS could be applied to maximise efficiency and simplicity. It is suggested here that the courts should always apply the 'but for' test first, since it is slightly simpler than NESS. Moreover, it would be unduly cumbersome to have to consider a simple case in terms of NESS. Only when 'but for' does not lead to a finding of causation should courts apply the NESS test. Since all 'but for' causes are NESSs, they would effectively be replacing the 'but for' test with a new test of factual causation, albeit in the simplest manner possible.

5. Conclusion

It is submitted that NESS is the most comprehensive single test of causation that the law could adopt and, therefore, would be an attractive replacement for the 'but for' test. Whereas the 'but for' test can be consistently applied and can limit liability in a meaningful way, it fails to meet the law's need for a fair test of factual causation. Conversely, such alternatives to the 'but for' test as the 'material contribution' test, cannot secure justice without sacrificing clarity in the law. The NESS test's practical utility lies in its ability to satisfy all three of the criteria identified in part II. While NESS' recognition of a broad range of causal involvement allows for fairer results, the reliance on the objective concepts of necessity and sufficiency ensures that the test can still limit liability in a clear-cut manner.

Admittedly, the test is not perfect. Part III identified various theoretical flaws, most notably NESS' inability, strictly speaking, to make a finding of causation in cases of indeterminacy. There are also practical dangers such as the 'proliferation of NESSs' or the fact that NESS' reliance on lawful sufficiency may make causation too easy to establish. Nevertheless, Part III demonstrated that the test can provide a solution to these problems. NESS *can*, in practice, rule out spurious causes such as the barometer 'causing' a storm. Used in a certain way, it *can* detect indeterminate causation. With that point in mind, although it has been argued that the test should be adopted into Scots law, it is argued here that the following supplementary rules would serve to maximise its utility: namely a rebuttable presumption that lawful sufficiency will be enough to establish causation¹³² and a 'no longer definitely sufficient' rule.¹³³

Moreover, it is fair to say that many of the practical dangers of introducing NESS are outweighed by the benefits. Fischer's 'tsunami' example¹³⁴ quite correctly demonstrates that NESS can lead to absurd conclusions, however if the occasional

¹³⁰ *ibid.*

¹³¹ These are all the forms of 'involvement' identified by Stapleton (n 3): see footnotes 85-87.

¹³² Discussed above: See text accompanying footnote 113.

¹³³ Discussed above: See text accompanying footnote 98.

¹³⁴ Fischer, (n 5) 'Insufficient Causes' 290-291.

bizarre finding of causation is the price of adopting a test that can recognise a wider category of causation in a logical, consistent manner, it is a price worth paying.

In justifying the 'increase in risk' rule in *McGhee*,¹³⁵ Lord Reid commented that the law did not need a philosophically laden account of causation, only one based on common sense.¹³⁶ If this article has demonstrated anything, however, it is that adopting a test of factual causation that is logical, rational and dogmatically rigid, can have considerable practical value. Indeed, it is rather puzzling that judges are so averse to 'big theory',¹³⁷ for having one test suitable for all causation problems would surely make the law clearer, simpler and more consistent. Given the practical benefits identified above, it seems that NESS, short of being adopted, should at least be taken seriously as a viable alternative to the current law. Judges may have perfectly valid reasons for dismissing NESS as unsuitable, but to ignore the test, to not even consider the potential benefits that would come of adopting it, would be remiss.

¹³⁵ *McGhee* (n 10).

¹³⁶ *ibid* 5.

¹³⁷ V Black, 'Decision Causation: Pandora's Tool-Box' in JW Neyers, E Chamberlain and SGA Pitel (eds), *Emerging Issues in Tort Law* (Hart Publishing 2007) 310.

Execution in Counterpart in Scots Law

CAROLINE M. HOOD*

Abstract

A specific aspect of English law which is inherently attractive to corporate and commercial practitioners in Scotland is that of 'execution in counterpart'. This refers to the process by which a contract may be signed, or executed, by each party signing its own copy - the counterpart - and then exchanging it with the other party or parties for their signed counterpart.¹ This paper will demonstrate that, in principle, Scots Law recognises the completion of a contract through an exchange of signed counterpart documents. Unlike the current position in English law, however, Scots law does not recognise execution in counterpart for the purposes of documents which require formal execution, due to the perceived lack of legal clarity on the self-proving status of such documents. Therefore, for the purposes of remote, all-party signings, reliance on Scots law is insufficient for the document to be granted self-proving status. The aim of this paper is to demonstrate that, in order to alter the current reliance on English law for the purposes of execution in counterpart, reform of the law is required and that '(...) the profession will need the assistance of the SLC [Scottish Law Commission] and academics to explain, in simple terms, the rationale and effect of the proposals, not only to the profession but also to their clients.'² This paper will also suggest that it remains crucial for Scots law to adopt an authoritative position on execution in counterpart to ensure certainty in the law of Scotland.

1. Introduction

As part of their Eighth Programme of Law Reform,³ the Scottish Law Commission recently published a Discussion Paper entitled *Review of Contract Law: Discussion*

* Graduate, School of Law, University of Aberdeen; Trainee Solicitor, Stronachs LLP. My thanks to Colin Mackie who provided helpful comments and suggestions on an earlier draft of this paper.

¹ Scottish Law Commission, *Review of Contract Law: Discussion Paper on Formation of Contract* (Discussion Paper No.154, March 2012) Glossary. (Hereafter, the 'Paper').

² Law Society of Scotland, 'Scottish Law Commission: Review of Contract Law Discussion Paper on Formation of Contract [Discussion Paper No. 154] The Law Society of Scotland's Response' 2-3.

³ The Programme states: 'We propose to review the law of contract in the light of the publication in 2009 of the Draft Common Frame of Reference: Principles, Definitions and Model Rules of European Private Law (the DCFR). The DCFR provides a contemporary statement of contract law, based on comparative research from across the European Union and written in accessible and non-archaic English. The DCFR has a considerable amount to offer in the law reform process. It may be seen as an instrument to provide an important area of Scots law with a systematic health check, giving a basis for treatment where the law is found to be ailing or otherwise in need of remedial treatment. The DCFR is at least a good working platform for a series of discrete and relatively limited projects on contract law, akin in some ways to our work on trusts and having significance for the well-being of the Scottish

*Paper on Formation of Contract.*⁴ It was during the course of this Paper that the Commission said the following:⁵

(...) it was suggested (...) that we should look at the question of “execution in counterpart” in contracts. We were told that this area was causing considerable difficulties for Scottish commercial practitioners in particular, to the extent that for some it was preferable to contract under English law as providing a better solution to the problems.

A practical example serves to illustrate the point. English law may be preferable to Scots law for the signing of a share purchase agreement, negotiated by Scottish solicitors on behalf of their clients, a company registered in Scotland. Due to logistical constraints, the signing parties will not be available to participate in an all-party signing and therefore a remote signing is proposed. The current position in Scots law is such that too much uncertainty surrounds whether a court would be willing to grant such a remotely signed document self-proving status. Therefore, in order to circumnavigate the difficulties of remote party signings in Scots law, the parties elect to insert a ‘choice of law’ clause to govern the contract by means of English law, thereby providing the certainty of a self-proving status for the executed contract. The clause inserted into the share purchase agreement would be similar to the following: ‘This Deed may be executed in any number of counterparts, each of which, when executed and delivered, shall constitute an original, but all the counterparts shall together constitute one and the same instrument’. The ‘choice of law’ clause ensures that English law governs the contract, meaning that execution in counterpart is possible.

The choice of English law in such circumstances is borne out of historical divergence in the law of contract between Scotland and England and Wales.⁶ The two jurisdictions have developed this area of law differently. In England, the dominance of the common law has allowed contract law to develop gradually through the courts. Indeed, this lack of ‘legislative interference’⁷ is credited with allowing English contract law to become a product of the ‘vigorous common sense of English judges’.⁸ The development of Scots contract law has been somewhat different:⁹

The Scottish system of contract law was developed (...) from the principles of the law of obligations that had been expounded by the jurists. The role of the judges was to fill in gaps where they were found and to develop and apply the basic principles. To some extent they could be creative in carrying out

economy.’ Scottish Law Commission, *Eighth Programme of Law Reform* (Scot Law Com No 220, 2010) [2.16].

⁴ Scottish Law Commission (n 1).

⁵ *ibid* [6.1].

⁶ For the purposes of clarity, all future references to England/English law should be deemed to include Wales.

⁷ Sir W Anson, *Principles of the English Law of Contract and of Agency in relation to Contract* (1st Edition, 1879) Preface; quoted in H Beale et al (eds), *Chitty on Contracts: Volume I: General Principles* (31st edn, Thomson Reuters 2012) [1-003].

⁸ *ibid*.

⁹ Lord Hope, ‘The Role of the Judge in Developing Contract Law’ (2011) 15 (1) Jersey and Guernsey Law Review (online edition) [20].

these functions. But their duty was to apply the law as they believed it to be. Their approach is, I think, inevitably, conservative rather than revolutionary. There are limits to the extent that the judges can reform the law. Structural changes must be left to the legislators.

The absence of a robust common law foundation for contract law in Scotland has led to recommendations for such 'structural changes' to legislation within the Scottish Law Commission's report, including recommendations in relation to execution in counterpart.

That there should be such a close connection between the UK jurisdictions on common matters of civil law is unsurprising, particularly given the commercial relationships between the two jurisdictions. However, it appears anomalous that such a fundamental aspect of the law in Scotland – contract law – is governed by English law in instances of remote party signings due to an on-going uncertainty in Scots law surrounding the status of documents executed in counterpart. The aim of this paper is to demonstrate that, in order to alter the current reliance on English law by Scots practitioners for the purposes of execution in counterpart, '(...) the profession will need the assistance of the SLC and academics to explain, in simple terms, the rationale and effect of the proposals, not only to the profession but also to their clients.'¹⁰ This aim will be achieved through analysis of the current position of execution in counterpart within Scots and English law and consideration of precisely what inspires Scots lawyers to elect to rely upon English contract law. This paper will also suggest that Scots law must adopt an authoritative position to safeguard certainty and the continued independence of the Scottish legal system.

This paper will be divided into three substantive sections. In the first section, the legal status of execution in counterpart in the context of English law will be considered in order to provide a basis for the analysis of Scots law. In the second, the position in Scots law will be analysed to establish an understanding of why there is uncertainty over the status of documents executed in counterpart. The third section will consider prospective developments which may enhance the certainty of the Scots law position. Finally, conclusions will be drawn as to whether these reforms will impact upon choice of law in corporate and commercial contracts, thereby allowing Scots law to become the choice of law in contracts which are sought to be executed in counterpart.

2. Execution in Counterpart in English Law

A. Background: Formation of Deeds in English Law

Under English law, '(...) contracts can be made quite informally, no writing or other form is necessary',¹¹ and, '(...) all formal requirements in the law of contract are contained in legislation which deals with specific contracts.'¹² However, it should be

¹⁰ Law Society of Scotland (n 2) 2-3.

¹¹ Beale et al (n 7) [4-001].

¹² *ibid.*

noted that varying requirements for specific forms of contracts lie within these formalities.¹³ In order to be accorded special status, a 'deed' must satisfy specific formal requirements. Put simply, a deed is '(...) a written instrument which is executed with the necessary formality'¹⁴ which confirms '(...) an interest, right, or property passes or is confirmed or an obligation binding on some person is created or confirmed.'¹⁵ In order to achieve this special status, therefore, a deed must be in writing and capable of satisfying certain requirements.

At common law, the only formal requirements for a deed were that it was in writing, sealed and subsequently delivered.¹⁶ However, as noted above, there are now statutory requirements in place, governing the valid execution of deeds.¹⁷ These requirements provide specific guidance to the format, parties, execution and delivery of the deed.¹⁸ It is these qualities which distinguish a deed from other ordinary documents such as a document entered into 'under hand'.¹⁹ While the distinction may seem arbitrary, it is an important one. There are two main differences between a deed and a document entered into under hand. First, a promise made without valuable consideration cannot generally be enforced. However, if the promise is made in a deed, this is known as a 'specialty' and the promise can usually²⁰ be enforced, despite the lack of consideration.²¹ The second difference relates to the period of limitation. For a general contract, the period of limitation extends to six years, whereas for a specialty, there is a twelve year period in which to make a claim.²² Another noteworthy difference between deeds and simple contracts is that there are further statutory provisions contained within the Companies Act 2006 for deeds executed for or on behalf of a company. Sections 44 to 47 specify circumstances under which a document will be validly executed as a deed by a company: (i) the company may affix its common seal; (ii) the document may be signed by two authorised signatories of the company; or (iii) a director may sign the document along with a witness who attests the signature.

Once the execution requirements are complied with, the document is then validly executed as a deed, granting the document self-proving status. This has the effect of creating an evidential presumption. Therefore, should the validity of the deed ever be questioned, the collective counterparts showing the signatures of the

¹³ *ibid* [4-002].

¹⁴ Law Commission, *The Execution of Deeds and Documents by or on behalf of Bodies Corporate* (Law Com No. 253) [2.4].

¹⁵ *ibid*.

¹⁶ *ibid* [2.8].

¹⁷ Law of Property (Miscellaneous Provisions) Act 1989. Section 1 specifically covers the circumstances under which a deed shall be deemed as having been validly executed.

¹⁸ *ibid* s.1(1), s.1(2)(a), s.1(2)(b) and s.1(3)(b) respectively.

¹⁹ 'The law of England and Wales makes an important distinction between instruments which are executed in "solemn form" as deeds, and other instruments which are generally referred to as being in "simple form" under hand. Where the instrument is a contract, this is recognised by the distinction made between a contract executed as a deed, which is a specialty, and an instrument under hand, which is a simple contract.' Law Commission (n 14) [2.2].

²⁰ Equity would not assist a volunteer.

²¹ Law Commission (n 14) [2.6].

²² *ibid*. The Law Commission also goes on to note in detail some further distinctions between deeds and other documents.

parties to the deed is sufficient proof of its validity and no further evidence is required. In sum, the certainty the document provides has both an evidential and commercial value understood by practitioners and their clients alike.

B. Execution in Counterpart

Following the case of *R (on the application of Mercury Tax Group and another) v HMRC*,²³ English case law on the matter of execution in counterpart was revisited and subsequently clarified. There, the court distinguished the earlier authoritative case of *Koenigsblatt v Sweet*,²⁴ by confirming that the transfer of signature pages from an incomplete version of a deed to a later, complete and amended version was not effective. This decision questioned the previous position in English law, traditionally demonstrated by the *Koenigsblatt* case, which held that where a contract was altered by an agent after signing, the subsequent ratification of the party was held to be lawful. This decision led to discussion about the status of pre-signed signature pages and the exchange of signature pages by email during the course of virtual signings and closings,²⁵ all means by which contracts were executed in counterparts to avoid the practical inconvenience of all-party signings.

As a result of the decision in *Mercury*, the Law Society of England and Wales issued a Practice Note with the underlying intention of clarifying the legal position regarding execution of documents at virtual signings or closings where some or all of the signatories are not physically present at the same meeting.²⁶ This was achieved as part of a Joint Working Party which analysed the relevant law and offered the following perspective on the *Mercury* case:²⁷

It is the view of Leading Counsel and of the JWP [Joint Working Party] that the Court of Appeal decision in *Koenigsblatt v Sweet* [1923] 2 Ch 314 (*Koenigsblatt*) remains the leading authority on the applicability of the principles of authority and ratification to the creation of legally binding written agreements, that *Mercury* (a first instance decision) should be viewed as limited to its particular facts and, to the extent inconsistent [sic] with *Koenigsblatt* (a Court of Appeal decision), the *Koenigsblatt* decision should prevail.

Therefore, despite the initial appearance that *Mercury* may have made substantial changes to the law in this field, it now appears that each case ought to be approached on its own facts when assessing which signing option is most appropriate.²⁸

²³ *R (on the application of Mercury Tax Group and another) v HMRC* [2008] EWHC 2721.

²⁴ *Koenigsblatt v Sweet* [1923] 2 Ch 314.

²⁵ Law Society, 'Practice Note: Execution of documents by virtual means'

<<http://www.lawsociety.org.uk/advice/practice-notes/virtual-execution-of-documents/>> [1.2] accessed 12 October 2012.

²⁶ *ibid.*

²⁷ Joint Working Party, 'Note on Execution of Documents at Virtual Signing or Closing' (May 2009; amended February 2010) <<http://www.citysolicitors.org.uk/FileServer.aspx?oID=760&IID=0>> accessed 15 October 2012.

²⁸ *ibid.*

Ultimately, the conclusion to the Joint Working Party's analysis of *Mercury* and *Koenigsblatt* came in the form of guidelines intended to prevent deeds from being incorrectly signed and therefore denied self-proving status. This clarification emphasises the need for certainty in such a crucial area of the law and draws into sharp contrast the current uncertainty in Scots law regarding execution in counterpart. The advice provided for practitioners in England and Wales is as follows:²⁹ (i) prior to signing, arrangements for the virtual signing are agreed between the parties' lawyers; (ii) final execution copies are then emailed to all absent parties; (iii) the signature page is printed and signed by each absent party; (iv) each absent party then returns a single email attaching the final version of the document and a pdf copy of the signed signature page. The parties can choose to state that signature of the document does not mean that the document has also been delivered; and (v) to evidence the execution of the final document, a final version of the document, together with copies of the executed signature pages, will then be circulated by one of the parties' lawyers. The final version of the document, complete with signed signature pages, will then constitute an original signed document, in line with the abovementioned requirements of the Law of Property (Miscellaneous Provisions) Act 1989 and Companies Act 2006 for execution of a document as a deed. Parties will then be provided with a certified copy of the deed and all the signed counterparts, which provides each party with proof that the contract has been validly executed. Should a question later arise as to the contract's validity, the individual contracts can each be re-constituted, demonstrating that they were all individually executed to create the final execution version of the contract.

In sum, execution in counterpart under English law is regarded as a competent method for the execution of deeds. In general terms, provided the requirements prescribed in the Law of Property (Miscellaneous Provisions) Act 1989 and Companies Act 2006 are complied with, a deed will be regarded as having been validly executed and will obtain self-proving status. However, following the *Mercury* case, the message for practitioners is that they should ensure that the guidelines of the Joint Working Party are observed. By doing so they will make certain that, in instances of remote signings, the validity of the executed deed cannot be later called into question. The certainty provided under English law, through both the jurisprudence of the courts and the statutory provisions mentioned above, is a distinct contrast to the current position in Scots law. Furthermore, this lack of certainty represents the root of the difficulty for practitioners in Scotland in the execution of deeds in counterparts.³⁰

²⁹ Adapted from *ibid* 4.

³⁰ See Journal of the Law Society of Scotland, 'Commission issues invite to 'execution' seminar' (15 November 2012) <<http://www.journalonline.co.uk/News/1011915.aspx>> accessed 16 November 2012.

3. Execution in Counterpart in Scots Law

A. Background: Formation of Contracts in Scots Law

Unlike English law, the word 'deed' has no specific meaning in Scots law. Nevertheless, there are certain characteristics attributed to documents purporting to be 'deeds', specifically: (i) formality; and (ii) an intention to create a legal relation.³¹ The lack of a statutory definition of a deed, however, has not proved problematic. The Inner House has been '(...) quite content to take 'deed' as being used in the popular sense',³² with Lord President Dunedin stating that: 'The statute has not defined what a deed is, nor am I tempted to give a definition'.³³ This position has been reiterated by the Court at subsequent points. First, in *Henderson's Trustees v Inland Revenue Commissioners*,³⁴ in considering the Stamp Act 1891, Lord President Dunedin comments:³⁵

(...) I think it unnecessary to consider whether the word 'deed' is there used as a term of art, because in any case it is only in England that it is so used. I am quite content to take 'deed' as being used in the popular sense. The statute has not defined what a deed is, nor am I tempted to give a definition, but I am certainly clearly of the opinion that, whatever is a deed, this acceptance of trust is not.

Furthermore, in the subsequent case of *Lennie v Lennie's Trustees*,³⁶ Lord Dewar states that:³⁷

In the absence of any definition of what a deed is, the question just comes to be whether it is reasonable to suppose that a business man would regard a pencil jotting on the back of an envelope as a 'deed'. I do not think that he would. (...) The word 'deed' would, I think, suggest to his mind some document his solicitors prepared, or at all events something of a much more formal character than we have here.

Therefore, as noted above, while the word 'deed' has no specific definition, there are certain characteristics attributed to documents purporting to be 'deeds', which must be achieved within the confines of Scots contract law.

There are specific statutory provisions in Scots law governing contract formation and, as in England, the provisions vary depending upon the nature of the contract. Practitioners and students will be acquainted with the Requirements of Writing (Scotland) Act 1995 (the '1995 Act'), which makes provisions for the

³¹ *Low Bonar plc v Mercer Ltd* [2010] CSOH 47, [16] (Lord Drummond Young).

³² *Henderson's Trustees v Inland Revenue Commissioners* 1913 SC 987 [989] (Lord President Dunedin).

³³ *ibid.*

³⁴ *Henderson's Trustees v Inland Revenue Commissioners* 1913 SC 987.

³⁵ *ibid* [989] (Lord President Dunedin).

³⁶ *Lennie v Lennie's Trustees* 1914 1 SLT 258.

³⁷ *ibid* [261] (Lord Dewar).

formalities of contractual documents. Whilst, in general,³⁸ writing is not required for a contract, in the absence of a written contract, the proviso *caveat emptor* applies and it therefore makes good commercial and practical sense for contracts to be in a written format. A written contract will provide an element of certainty with respect to the terms upon which the transaction is intended to proceed and the particulars of the contractual agreement. Section 48 of the Companies Act 2006 specifies that for a company to validly execute a deed under Scots law, the provisions of the 1995 Act must be complied with. Thus, while English law sets out the specific requirements for the execution of documents by companies, Scots law merely requires adherence to the existing legislation.

The 1995 Act is not the only statute governing Scots contract law, however. Where a contract is reduced to a written document, section 1(1) of the Contracts (Scotland) Act 1997 contains a presumption that the contract contains all the express terms of the contract. The written contract is granted validity through the act of signing, provisions for which are made in section 2 of the 1995 Act and, if the parties choose to have their signatures witnessed, then the document achieves self-proving status.³⁹ This presumption is difficult to defeat as, by implication, to do so would require proof that the signature is fraudulent. Therefore, by committing a contract to writing and having it witnessed, the self-proving status of the document grants both an evidential and commercial value. As a result, the parties have the dual benefit of both clarity of terms and the legal certainty of a written, legally enforceable document.

Finally, the matter of delivery of the contract ought to be considered. This position is best explained by the famous case of *Stamfield's Creditors v Scot's Children*⁴⁰ in which a granter of an assignation was found murdered with the signed document beside him. The court held that, as delivery had not been achieved, the assignation was ineffective.⁴¹ However, in the case of mutual contracts reduced to writing in a single document and signed by all the parties, there is no physical requirement for delivery of the document.⁴² Hence, the delivery requirement impacts most significantly upon unilateral obligations and mutual obligations which are set out in multiple copies, signed by each granter and exchanged between parties.⁴³

³⁸ There are certain exceptions. For example, writing will be required for the constitution of a contract to the creation, transfer, variation or extinction of an interest in land; gratuitous unilateral obligations (except those undertaken in the course of a business); the constitution of a trust where an individual declares himself to be sole trustee of his property or any property that he may acquire; and for the making of any will, testamentary trust disposition and settlement or codicil. These exceptions are provided for by section 1(2), Requirements of Writing (Scotland) Act 1995.

³⁹ In effect, sections 3(1) and (8) of the 1995 Act mean that if a signature is witnessed, it is presumed that the person who bears to have signed the document did so and also that it was signed on the date specified in the contract.

⁴⁰ *Stamfield's Creditors v Scot's Children* (1696) Br Supp IV 344.

⁴¹ *ibid*, quoted in *Park, Petitioners* [2009] CSOH 122 [10] (Temporary Judge MG Thomson QC).

⁴² See Scottish Law Commission (n 1) [7.4].

⁴³ *ibid*.

B. Execution in Counterpart

From the previous section it can be seen that the Scots approach to contract law, specifically the execution of documents, differs from the approach in England and Wales. This subtle but important difference has not gone unnoticed by the Scottish Law Commission who, in their Discussion Paper, commented with regard to execution in counterpart that, 'We were told that this area was causing considerable difficulties for Scottish Commercial practitioners (...) to the extent that for some it was preferable to contract under English law (...)'.⁴⁴ This concern is not merely borne out of a desire to ensure that fees are not lost to larger firms with Scottish and English offices, it also has two important practical implications: (1) for the future of Scots law as an independent legal system; and (2) if contracts with 'choice of law' provisions are drafted by Scots lawyers with little or no English law background, then drafting and legal errors could be made as certain consequences under English law may not have been foreseen. Furthermore, given the traditional independence of the Scots legal system, it seems paradoxical that Scots lawyers may seek to obtain further qualifications in English law to enable them to authoritatively draft a contract while sitting in their Scottish office, acting for their Scottish clients all in an effort to achieve certainty in the law and in execution. Critically, it is important that a 'choice of law' clause should not be seen as an 'easy way out' of a difficult situation; it is an exploitation of a loophole which may have unforeseen consequences.

The crux of the difficulty in Scots law regarding execution in counterpart can be summarised as follows:⁴⁵

(...) the problem of "execution in counterpart" arises when the parties are in different locations, albeit each with access to a print-out of the contract, and able to subscribe that copy and to transmit it electronically (by fax or email attachment of a scanned version) to emerge at the other end as a facsimile of the signed copy. Will this be enough for the parties to have a contract and, if the contract so provides, for each party's set of documents to have full contractual force?

The starting point for analysis of the uncertainty in Scots law is the common law. Two cases exist which support the position that, in principle, Scots law recognises an exchange of counterpart documents, each signed by its respective party, as a contract. Critically, however, neither case makes reference to the matter of delivery of the document. It is precisely this lack of guidance in case law which creates a legal difficulty in circumstances where parties are signing the documents remotely.

The 18th century case of *Smith v Duke of Gordon*⁴⁶ concerned circumstances in which, following a contract executed in counterpart, the pursuer obtained a decree against the Duke for a sum of money in respect of a contract which had not been fulfilled. The court held that:⁴⁷

⁴⁴ *ibid* [6.1].

⁴⁵ *ibid* [6.22].

⁴⁶ *Smith v Duke of Gordon* (1701) Mor 16987.

⁴⁷ *ibid* quoted by the court in *Wilson v Fenton Bros (Glasgow) Ltd* 1957 SLT (Sh Ct) 3 [5].

(...) mutual contracts having two doubles need not be subscribed by both parties-contractors, but it was sufficient in law if the Duke's principal was signed by Smith and his counterpart by the Duke; and it was so found lately, in a case of Sinclair of Ossory in Caithness; and therefore sustained the marginal note, though not signed by the Duke, seeing it was contained in his own double uncanceled". The marginal note or rubric in the report of the case reads: "If a mutual contract is executed by two counterparts, it is sufficient if each party subscribes the paper containing what is prestable on himself".

This case was cited as authority in the more recent case of *Wilson v Fenton Bros (Glasgow) Ltd*,⁴⁸ where a licence agreement was held to be validly executed despite the parties having executed separate copies. Here, the owners of a patent granted a licence for the manufacture and sale of the patented item to a limited company for an annual minimum payment. The licence was prepared in duplicate; one copy was executed by the owners and handed to the company, and the other was executed by the company and handed to the owners. In handing down his judgment, the sheriff-substitute stated that:⁴⁹

In my view, the documents produced and to which I have referred, establish the fact of a completed agreement between the pursuers and the defenders of date 23rd December 1948 (...) The form of agreement is not a usual one in Scotland, but, as all the negotiations were conducted in England, the method of having two copies, of which one copy is signed by each party and delivered to the other party, was adopted in conformity, as I am informed, with a common practice in England.

However, there appears to be no further authority in Scots law for this proposition,⁵⁰ particularly in the context of the execution of deeds. This inadequacy in Scots law results in commercial practitioners seeking to draft deeds subject to English law, with a view to obtaining certainty regarding execution in counterpart. The lack of clarity regarding execution in counterpart and remote party signings means that, under Scots law, a signing meeting will be held which all parties to the contract will attend. The document is then signed and witnessed in the presence of all parties and is thereby granted self-proving status.

The Discussion Paper and the Scottish Law Commission conclude that while the judicial recognitions of execution in counterpart in *Wilson v Fenton Bros (Glasgow) Ltd* and *Smith v Duke of Gordon* confirm the principle that such a process can form a contract in Scots law, neither case explicitly discusses the application in relation to delivery of documents or the self-proving status of such documents.⁵¹ They therefore conclude that while purely hard copy and physical delivery transactions may be possible under Scots law, electronic transmission of documents which are reproduced at the recipients' end in electronic hard copy must fail to satisfy the

⁴⁸ *Wilson v Fenton Bros (Glasgow) Ltd* 1957 SLT (Sh Ct) 3.

⁴⁹ *ibid* [5].

⁵⁰ No further authority is cited by the Scottish Law Commission and the author's own research did not expose any further cases.

⁵¹ Scottish Law Commission (n 1) [6.28].

current legal test for a document to achieve self-proving status.⁵² From the previously cited example of the share purchase agreement, it is apparent that Scots law is currently unsatisfactory and, for the purposes of simplicity, practitioners will elect to utilise a 'choice of law' clause and for English law to govern a contract. Therefore, in order to maintain the independence of the Scots legal system it is crucial that execution in counterpart is clearly permitted by Scots law; a position which must be achieved through considered reform of the existing legislation.

4. Conclusion: Clarifying the Law

Despite the evident uncertainty within Scots law regarding execution in counterpart, there are some pending statutory changes to the law in Scotland which may positively impact upon the situation. Specifically, when fully enacted, Part 10 of the Land Registration etc. (Scotland) Act 2012 will make significant modifications to the Requirements of Writing (Scotland) Act 1995. Crucially, section 96 of the 2012 Act will modify section 1 of the 1995 Act in relation to electronic documents. This will create a situation where, if a document is required to be in writing, the form of that document can be either a traditional document or an electronic document, so long as it complies with the specified regulations. Although the Act can be criticised for a lack of clarity on this point, the regulations are incorporated through a new Part III to the 1995 Act⁵³ and can potentially improve legal certainty in relation to electronic documents and, therefore, the capacity for contracts governed by Scots law to be executed in counterparts. Yet, despite this, it is suggested that there may still be a lack of certainty regarding electronic documents in instances where the relevant party prints a hard copy of the signing page and signs it by hand, rather than using an electronic signature. Such circumstances may be sufficient to create enough doubt in the mind of the court as to the self-proving status of such a contract that practitioners may still be unwilling to run the risk of a court failing to uphold the validity of a contract on this basis.

The 2012 Act is not the end of the reforms but merely the beginning of a series of developments to clarify Scots contract law. Indeed, the Scottish Law Commission was keen to emphasise both the format that further developments in Scots law should take, and that the E-Commerce Directive⁵⁴ ought to form a role in any prospective legal development.⁵⁵ Of particular relevance is Article 9(1) of the Directive which states that:⁵⁶

Member States shall ensure that their legal system allows contracts to be concluded by electronic means (...) ensure that the legal requirements

⁵² *ibid* [6.29].

⁵³ Inserted by section 97 of the Land Registration etc. (Scotland) Act 2012.

⁵⁴ Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market (Directive on electronic commerce) (OJL 178/1, 17.7.2000). This directive was incorporated into United Kingdom domestic law by the Electronic Commerce (EC Directive) Regulations 2002.

⁵⁵ Scottish Law Commission (n 1) [6.3].

⁵⁶ *ibid*; Directive 2000/31/EC (n 54).

applicable to the contractual process neither create obstacles for the use of electronic contracts nor result in such contracts being deprived of legal effectiveness and validity on account of their having been made by electronic means.

Therefore there is an obligation to ‘(...) ensure the free movement of “information society services” across the European Community and to encourage greater use of e-commerce’.⁵⁷ With specific reference to the conclusion of contracts by electronic methods, the UK Government states:⁵⁸

The Government believes that the great majority of relevant statutory references (e.g. to requirements for writing or signature) are already capable of being fulfilled by electronic communications where the context in which they appear does not indicate to the contrary. Where existing legal requirements applicable to the contractual process do create obstacles for the use of the electronic contracts or result in such contracts being deprived of legal effectiveness and validity on account of their having been made by electronic means, the Government will propose necessary amendments on a case-by-case basis.

The Directive creates an obligation upon the UK and Scottish Governments to facilitate the use of e-commerce and, in particular, the use of e-commerce in contracts. In their discussion paper, the Commission propose drafting a clear legislative provision stating that an exchange of counterpart contractual documents can satisfy the requirements of formal validity and probativity if each document is appropriately executed by the respective parties to the exchange, where the documents are electronically delivered and where there is also agreement to hold the originals as delivered to the other party/parties involved.⁵⁹ This would go some way toward clarifying the legal position and allowing for the conclusion of contracts by electronic means to achieve legal certainty in Scots law.

However, despite the Commission’s enthusiasm for future reform the Law Society of Scotland has urged caution over any proposed reforms to the law in this area. In a response to the Commission’s Discussion Paper the Law Society of Scotland stated that:⁶⁰

We can see that bringing Scots law into line with modern international benchmarks would, in principle, make Scots law easier to explain to a whole host of international clients. In practice, however, the horizons of in-house counsel are often limited to the English approach. In other words, if the SLC [Scottish Law Commission] is to proceed with these interesting proposals, the profession will need the assistance of the SLC and academics to explain, in

⁵⁷ Department for Culture, Media and Sport, ‘The Electronic Commerce Directive (00/31/EC) and the Electronic Commerce (EC Directive) Regulations 2002 (SI 2002 No.2013)’ (2008), 3
<<http://www.culture.gov.uk/images/publications/eCommerce-guidance.pdf>> accessed 15 October 2012.

⁵⁸ *ibid* 7.

⁵⁹ Scottish Law Commission (n 1) [7.28].

⁶⁰ Law Society of Scotland (n 2).

simple terms, the rationale and effect of the proposals, not only to the profession but also to their clients.

The Law Society of Scotland's position appears to be that they accept the Scottish profession's reliance upon English law and, moreover, that persuasive evidence would be required to convince solicitors to rely on Scots law to execute documents in counterpart, even after further reform.

Effective law reform is not achieved solely through statutory means but with, '[t]he contribution of politicians, judges, the legal profession, charities, pressure groups and individuals who sometimes effectively highlight an injustice'.⁶¹ To ensure the continuing relevance of Scots law in commercial transactions, there is a need for a consistent and coherent programme of law reform in connection with Scots contract law. Within such reform lies the opportunity to clarify and consolidate the position of execution in counterpart in Scots law and to generate certainty in an area of law marred by confusion. However, in order for the reforms to be truly successful, the rationale and applicability of any such reforms must be accepted by the legal profession who, in acceding to such change, must also ensure that they are acting in the best interests of their clients. Although execution in counterpart is not the sole reason practitioners will elect to insert a 'choice of law' clause in favour of English law into a contract, such as in the share purchase example cited above, the current certainty in English law regarding the self-proving status of documents executed in counterpart means that in instances where an all-party signing meeting is not possible, English law is the natural choice. In absence of persuasive evidence that Scots law accepts the self-proving status of documents executed in counterpart Scots law will, in such circumstances, remain subordinate to English law, at least in this aspect of commercial transactions, in the eyes of both practitioners and their clients.

⁶¹ Lady Clark of Calton, 'Lecture by Lady Clark of Calton as Chairman of the Scottish Law Commission by invitation of SCOLAG on 26 September 2012' (Scottish Law Commission, October 2012) <<http://www.scotlawcom.gov.uk/news/lecture-by-lady-clark-of-calton/>> accessed 2 November 2012.

Pragmatism in the European Union: Recasting the Brussels I Regulation to Ensure the Effectiveness of Exclusive Choice-of-Court Agreements

JUSTIN P. COOK*

Abstract

The Brussels regime of international instruments governing jurisdiction and the recognition and enforcement of judgments in civil and commercial matters provides a comprehensive framework of uniform rules for allocating jurisdiction and promoting the free circulation of judgments. In general, this regime has successfully facilitated cross-border litigation in the European judicial area. Nevertheless, it has been significantly criticised for its treatment of exclusive choice-of-court agreements. The European legislature intends to address such criticism with the recasting of the Brussels I Regulation. This article examines, inter alia, the key modifications in the Recast Regulation concerning this area, namely, the reversal of the priority ruling and the introduction of a harmonised choice-of-law rule. These reforms are purportedly designed to achieve two fundamental goals: to enhance the effectiveness of choice-of-court agreements and to ensure that such agreements are treated consistently in both the Recast Regulation and the Hague Convention of Choice of Court Agreements. It is suggested that, in the main, the proposed amendments should satisfy these objectives. However, the article also maintains that the reforms should have been extended to improve the treatment of jurisdictional agreements in favour of non-European States.

1. Introduction

The aim of the Brussels I Regulation¹ is to further the ideal of an internal market by promoting the free circulation of civil and commercial judgments within the European Union (EU).² This goal is achieved, in part, by strictly allocating jurisdiction in accordance with a hierarchy of jurisdictional rules.³ In 2009, however, a report on the application of the Regulation published by the European

* The author is an LLM candidate in Private International Law at the University of Aberdeen. The author would like to express his gratitude to the managing editor of the Aberdeen Student Law Review, Ilona Cairns, and the anonymous reviewer for their valuable comments and input.

¹ Council Regulation (EC) 44/2001 of 22 December 2000 on jurisdiction and the regulation and enforcement of judgments in civil and commercial matters, [2001] OJ L21/1 (the 'Regulation'). Hereinafter, the 'Brussels regime' shall refer to both the 1968 Brussels Convention on Jurisdiction and Judgments (the 'Brussels Convention') and the Regulation.

² The Regulation, Recital 6. See also A Bell, *Forum-Shopping and Venue in Transnational Litigation* (OUP 2003) 52.

³ J Steinle and E Vasiliades, 'The Enforcement of Jurisdiction Agreements Under the Brussels I Regulation: Reconsidering the Principle of Party Autonomy' (2010) JPIL 565, 587.

Commission⁴ identified *inter alia* two major concerns regarding the instrument's treatment of valid choice-of-court agreements.⁵ The first was that the regime inadequately protected exclusive choice-of-court agreements,⁶ while the second concern arose from the need to ensure that choice-of-court agreements were effective both within and beyond EU borders.⁷ The Report also highlighted the need for coherency between the rules of the Recast Regulation and the Hague Choice of Court Agreement Convention (HCCA),⁸ a treaty which primarily focuses on improving the legal certainty of jurisdiction agreements in international commercial relationships.⁹

In December 2010, the European Commission advanced the process for reforming the Regulation by releasing its Recast Proposal.¹⁰ The draft amendments proposed by the Commission were subjected to an extensive review by the European legislature and it was not until 20 November 2012 that the European Parliament voted in favour of the revised text.¹¹ On 6 December 2012, the EU Council also

⁴ European Commission, 'Report from the Commission to the European Parliament, the Council and the European Economic and Social Committee on the application of Council Regulation (EC) No 44/2001 COM (2009) 174 final [21.4.2009]. Hereinafter this report shall be referred to as the 'Commission's Report' or the 'Report'. See also Article 73 of the Regulation, which provides that, no later than five years after its entry into force, the Commission shall prepare a report on the operation of the Regulation.

⁵ *ibid* (COM (2009) 174 final) 5-6. See also European Commission Staff Working Paper, 'Impact Assessment' SEC (2010) 1547 final, 28-9.

⁶ *ibid* (COM (2009) 174 final) 6. An exclusive jurisdiction clause not only confers jurisdiction on a particular court but also precludes a party from litigating in another forum. Alternatively, a non-exclusive jurisdiction clause merely amounts to a submission to the jurisdiction of the specified court and does not prevent litigation from being initiated elsewhere. See J Fawcett, 'Non-exclusive Jurisdiction Agreements in Private International Law' [2001] LMCLQ 234, 234-36. Note that jurisdiction agreements in both the Regulation and the Hague Convention of 30 June 2005 on Choice-of-Court Agreements (HCCA) are presumed to be exclusive. Article 23(1) of the Regulation provides that: '[j]urisdiction shall be exclusive unless the parties have agreed otherwise', while Article 1 of the HCCA holds that this treaty applies in '(...) international cases to exclusive choice-of-court agreements concluded in civil and commercial matters.' References to choice-of-court agreements in this article should be assumed to be of the exclusive form. Hereinafter, the HCCA shall also be referred to as 'the Convention'.

⁷ *ibid* (COM (2009) 174 final) 6.

⁸ *ibid*.

⁹ Impact Assessment (n 5) 29. The HCCA is not yet in force. However, on 26 September 2007, Mexico became the first State to accede to the Convention. The United States signed on 19 January 2009 and the European Union followed on 1 April 2009. To date, Mexico, the United States and the European Union are the only States to have signed the Convention. See Hague Conference on Private International Law, 'The Hague Convention of 30 June 2005 on Choice of Courts Agreements' <http://www.hcch.net/index_en.php?act=conventions.status&cid=98> accessed 29 June 2013.

¹⁰ European Commission, 'Proposal for a Regulation of the European Parliament and of the Council on jurisdiction and the recognition and enforcement of judgements in civil and commercial matters (Recast)' COM (2010) 748 final [14.12.2010]. Hereinafter this report shall be referred to as the 'Commission's Proposal' or the 'Proposal'.

¹¹ European Parliament, 'Legislative resolution of 20 November 2012 on the proposal for a regulation of the European Parliament and of the Council on jurisdiction and the recognition and enforcement of judgements in civil and commercial matters (recast)' (COM(2010) 0748 - C7-0433/2010 - 2010/0383(COD))

<<http://www.europarl.europa.eu/sides/getDoc.do?type=TA&language=EN&reference=P7-TA-2012-412#BKMD-1>> accessed 10 June 2013. See also L Timmer, 'Abolition of Exequatur under the Brussels I Regulation: Ill conceived and Premature?' (2013) JPIL 129, 129.

adopted the Recast.¹² It is worth noting that, although the instrument entered into force on 9 January 2013, the revised European jurisdictional regime will apply only to proceedings initiated on or after 10 January 2015.¹³

The Recast incorporates a number of measures that purportedly enhance the effectiveness of choice-of-court agreements. The fact that the Regulation has been described as '(...) the most successful instrument on international civil procedure of all time'¹⁴ and is considered a fundamental component of commercial and legal practice in the EU requires that the revisions to the Regulation be 'closely scrutinised' and 'not accepted without demur'.¹⁵ In light of these strong claims, and the significance of the Brussels regime to the operation of the internal market, the central purpose of this article is to critically assess the amendments in the Recast.

This article will initially elaborate on the main issues that obscured the enforcement of jurisdiction agreements under the Regulation. Following an outline of the relevant amendments in the Recast, a critical evaluation of the key reforms to the law governing choice-of-court agreements in Europe shall be presented. The reforms will then be assessed in the context of the HCCA. The article maintains that the proposed amendments are not only necessary but, in view of the growth in cross-border business transactions in the global economy and the concomitant demand for greater certainty in international commercial litigation, can rightfully be described as pragmatic and meaningful. Firstly, however, it is essential to appreciate the problems that currently taint the relationship between the Regulation and choice-of-court agreements.

2. *The Interaction Between Choice-of-Court Agreements and the Regulation*

The parties' autonomous choice of venue is accorded significant respect within the Regulation's rigid framework of jurisdictional rules.¹⁶ For instance, Article 23 of the Regulation provides for the exclusive jurisdiction of a Member State's courts where the parties have incorporated an exclusive choice-of-court agreement prorogating

¹² European Council, 'Recast of the Brussels I Regulation: towards easier and faster circulation of judgments in civil and commercial matters within the EU' (6 December 2012) <http://europa.eu/rapid/press-release_PRES-12-483_en.htm> accessed 10 June 2013.

¹³ E Crawford and J Carruthers, 'Brussels I bis - the Brussels Regulation recast: closure (for the foreseeable future)' (2013) SLT 89, 89. For the adopted text, see Regulation 1215/2012 of the European Parliament and of the Council of 12 December 2012 on jurisdiction and the recognition and enforcement of judgements in civil and commercial matters (recast), [2012] OJ L351/1. Hereinafter the adopted recast Regulation shall be referred to as the 'Recast Regulation' or the 'Recast'.

¹⁴ R Goode, H Kronke and E McKendrick, *Transnational Commercial Law - Primary Materials* (OUP, 2007) 793 cited in B Hess, T Pfeiffer and P Schlosser, 'Report on the Application of Brussels I Regulation in the Member States (Study JLS/C4/2005/03)' (September, 2007) <http://ec.europa.eu/civiljustice/news/docs/study_application_brussels_1_en.pdf> accessed 10 June 2013.

¹⁵ A Dickinson, 'The Proposed "Brussels I Bis" Regulation' (*European Parliament, Directorate-General for Internal Policies, Legal Affairs*, 2011) <www.papers.ssrn.com/sol3/papers.cfm?abstract_id=1930712&download=yes> accessed 25 November 2012.

¹⁶ R Fentiman, 'Case Note on *Erich Gasser GmbH v Misat Srl*' 42 CMLR 241, 242. See also Case C-116/02 *Erich Gasser GmbH v Misat Srl* [2003] ECR I-14693[30] (hereinafter '*Gasser*').

that particular court.¹⁷ This provision is an exception to the general rule contained in Article 2 which specifies that the defendant can only be sued in their country of domicile.¹⁸ The Brussels regime, though, must also consider other values and a fundamental objective is to negate the risk of producing irreconcilable judgments.¹⁹ The possibility of concurrent proceedings, and subsequent inconsistent decisions, is prevented by a 'simple test of chronological priority'²⁰ articulated in Article 27(1) of the Regulation that grants precedence to the first seised court in resolving conflicts of jurisdiction.²¹ This provision provides that:

[w]here proceedings involving the same cause of action between the same parties are brought in the courts of different member States, any court other than the court first seised shall (...) stay its proceedings until such time as the jurisdiction of the court first seised is established.

The test expressed in the Regulation's Article 27 is also referred to as the 'first-in-time' rule or the doctrine of *lis pendens*.

Gardella and Radicati di Brozolo maintain that the *lis pendens* doctrine is designed to ensure 'neutrality, predictability and certainty' in litigation.²² Further, these authors contend that this procedural system for allocating jurisdiction is typical of the civil law tradition which has largely influenced the rules of the Brussels regime.²³ In effect, the first-in-time rule creates certainty by strictly curtailing the discretionary power of the judiciary. This approach contrasts starkly with the common law doctrine of *forum non conveniens* for resolving conflicts of jurisdiction, which effectively proposes that a court may decline jurisdiction on the grounds that it is not the appropriate forum and that considerations of justice require that the dispute is resolved in the courts of another State.²⁴

Critically, the Court of Justice of the European Union (CJEU) unequivocally endorsed the *lis pendens* doctrine in *Gasser*, the leading authority concerning the interplay between Articles 23 and 27 of the Regulation. In this case, an Austrian manufacturer and supplier, Erich Gasser, and MISAT, an Italian distributor, held a long-term contract for the supply of children's clothing.²⁵ The contract included a clause that granted exclusive jurisdiction to the courts of Austria.²⁶ MISAT, in breach of the jurisdictional clause, commenced proceedings before a court in Rome for a

¹⁷ Sir A Clarke, 'The Differing Approach to Commercial Litigation in the European Court of Justice and the Courts of England and Wales' IALS conference, (London February 2006) <www.dca.gov.uk/judicial/speeches/2006/sp060223.htm> accessed 2 November 2012.

¹⁸ The Regulation, Recital 14. See also J Harris, 'Understanding the English response to the Europeanisation of Private International Law' (2008) 4 JPIL 347, 352.

¹⁹ P Briza, 'Choice-of-Court Agreements: Could the Hague Choice-of-Court Agreements Convention and the Reform of the Brussels I Regulation be the way out of the *Gasser-Owusu* Disillusion?' (2009) 5 JPIL 537, 538.

²⁰ J Mance, 'Exclusive Jurisdiction Agreements and European Ideals' (2004) 120 LQR 357, 358.

²¹ Clarke (n 17) 7.

²² A Gardella and L Radicati di Brozolo, 'Civil Law, Common Law and Market Integration: The EC Approach to Conflicts of Jurisdiction' (2003) 51 Am J Comp L 611, 615.

²³ *ibid* 612.

²⁴ *Spiliada Maritime Corp v Cansulex Ltd* [1987] AC 460, [476] (Lord Goff).

²⁵ *Gasser* (n 16) [11].

²⁶ *ibid* [12].

declaration that the contract had been terminated. The supplier subsequently brought proceedings before the designated Austrian court.²⁷ MISAT contested the validity of the Austrian court's jurisdiction claiming that, pursuant to Article 21 of the Brussels Convention (now Article 27 of the Regulation), it was for the Italian court to decide this matter as the first seised court.²⁸ The primary question referred by the Austrian appellate court to the CJEU for clarification was whether the second seised court, which had exclusive jurisdiction pursuant to the choice-of-court agreement, could review the jurisdiction of the first seised court.²⁹ The CJEU confirmed that if one party commences proceedings in breach of a jurisdiction agreement before the other party brings an action in the putatively chosen court, the second seised court must stay proceedings until the first court has ascertained whether it holds jurisdiction.³⁰

Gasser has been subjected to substantial criticism. Although the Court upheld the strictures of the first-in-time rule,³¹ legal commentators objected to the fact that the decision encourages 'abusive litigation tactics'³² in commercial practice as proceedings may be initiated in 'bad faith' in a forum other than that agreed upon by the parties.³³ A race to court often eventuates to effect control of the forum.³⁴ Briggs asserts that this outcome is 'lamentable' and attributes the blame on '(...) the weakness of Article 23, the stiffness of Article 27 and *Gasser*.'³⁵ Furthermore, in *Turner v Grovit*,³⁶ the CJEU ruled that the use of anti-suit injunctions, which can be used to strengthen the effect of jurisdiction agreements by restraining a party from pursuing proceedings before a non-chosen court,³⁷ is incompatible with the principle of mutual trust that the Member States are required to extend to '(...) one another's legal systems and judicial institutions.'³⁸

The CJEU's case law has been regularly criticised for promoting tactical manoeuvres that are commonly referred to as 'Italian torpedo' actions.³⁹ This is a particularly effective tactic employed by unscrupulous parties whereby the court of a Member State with an inordinately slow procedural system is seised first.⁴⁰ The subsequent delay can secure an advantage for the party that launched the initial proceedings in the non-designated or 'wrong' court by frustrating the opposing

²⁷ *ibid* [13].

²⁸ *ibid* [14].

²⁹ *ibid* [20].

³⁰ *ibid* [54].

³¹ Harris (n 18) 368.

³² Impact Assessment (n 5) 28.

³³ Harris (n 18) 368. See also D Sancho Villa, 'Jurisdiction over Jurisdiction and Choice-of-Court Agreements: Views on the Hague Convention of 2005 and Implications for the European Regime' (2010) 12 Yrbk Priv Intl L 399, 402.

³⁴ *ibid* (Sancho Villa) 402.

³⁵ A Briggs, 'What Should Be Done about Jurisdiction Agreements?' (2010) 12 Yrbk Priv Intl L 311, 316.

³⁶ *Turner v Grovit* [2004] ECR I-3565.

³⁷ P Cachia, 'Recent Developments in the Sphere of Jurisdiction in Civil and Commercial Matters' (2011) *Elsa Malta Law Review* 69, 77.

³⁸ *Turner v Grovit* (n 36) [24].

³⁹ Cachia (n 37) 77.

⁴⁰ T Hartley, *International Commercial Litigation* (CUP 2009) 254-5.

party in terms of both time and the extra costs incurred from protracted litigation.⁴¹ Mance asserts, therefore, that the CJEU's literal interpretation of the Regulation presents real 'problems for legitimate claimants' and opportunities for those litigants who are unwilling to meet their contractual obligations.⁴²

In addition, the absolutist interpretative approach adopted by the CJEU has significant potential to undermine the effectiveness of jurisdiction agreements which nominate the courts of a non-Member State.⁴³ On this point, there is no provision in the Regulation's text that directly regulates the relationship between the courts of Member States and non-Member States.⁴⁴ Nevertheless, the CJEU's decision in *Owusu v Jackson*⁴⁵ has caused 'general unease' amongst a number of common law lawyers on the basis that the Court's jurisprudence does not adequately protect the parties' intention to litigate in a non-EU court.⁴⁶

In *Owusu*, the plaintiff, a UK domiciliary, was seriously injured while on vacation in Jamaica. Mr Owusu brought suit in the High Court of England and Wales against Mr Jackson, the owner of the premises where the plaintiff was injured, and five Jamaican legal entities that supervised the premises.⁴⁷ Crucially, Jackson was also domiciled in the UK. The defendants sought for the proceedings to be stayed, basing their claim on the common law doctrine of *forum non conveniens* and arguing that Jamaica was the natural forum.⁴⁸ The CJEU rejected the application and ruled that the mandatory nature of the jurisdictional rule in Article 2 prohibited a stay of proceedings under the principle of *forum non conveniens*.⁴⁹ The concern of some UK conflicts scholars with respect to *Owusu* is that the mandatory acquisition of jurisdiction under the auspice of Article 2 could override a mutual agreement that confers jurisdiction on a non-EU court.⁵⁰ It is suggested that the Court's commitment to legal certainty in this case highlights the ideological divide between the civil law and common law traditions over jurisdictional issues.⁵¹

The applicable law governing the validity of the parties' agreement is another source of uncertainty within the Regulation.⁵² The CJEU provides that in the context of validity, the concept of an 'agreement conferring jurisdiction' must be construed autonomously.⁵³ Essentially, the case law illustrates that if the formal requirements

⁴¹ T Hartley, 'The European Union and the Systematic Dismantling of the Common Law of Conflict of Laws' (2005) 54 Int'l & Comp LQ 813, 816-7.

⁴² Mance (n 20) 357.

⁴³ Harris (n 18) 371. See also Mance (n 20) 362.

⁴⁴ CMV Clarkson and J Hill, *The Conflict of Laws* (3rd edn, OUP 2006) 117.

⁴⁵ *Owusu v Jackson* Case C-281/02 [2005] ECR I-1383 (hereinafter '*Owusu*').

⁴⁶ Harris (n 18) 371. See also E Peel, 'Forum non conveniens and European Ideals' [2005] LMCLQ 363, 367-72.

⁴⁷ *Owusu* (n 45) [11], [12] and [14].

⁴⁸ *ibid* [15].

⁴⁹ *ibid* [46].

⁵⁰ A Briggs, *Agreements on Jurisdiction and Choice-of Law* (2008 OUP) 294.

⁵¹ Harris (n 18) 381.

⁵² Sancho Villa (n 33) 402. Such disputes generally relate to the 'existence of the agreement' - for example, whether there was proper consent - and other elements regarding the substantive validity of the impugned agreement such as capacity, fraud or duress (see Sancho Villa (n 33) 400).

⁵³ Case 214/89 *Powell Duffryn plc v Petereit* [1992] ECR I-1745, [14]. See also A Dickinson, 'Surveying the Proposed Brussels I bis Regulation' (2010) 12 Yrbk Priv Intl L 247, 285.

are satisfied,⁵⁴ then consensus is established for the purpose of upholding the choice-of-court agreement.⁵⁵ Under this interpretation, there is 'little room for national law'.⁵⁶ That construction, though, is open to criticism because '(...) if the national court is not entitled to examine the essential validity of the agreement' the fundamental purpose of Article 23 may be frustrated.⁵⁷

Moreover, the Commission's Report highlighted that not only is the substantive validity of jurisdiction agreements sometimes regulated on a residual basis by the various national laws but difficulties can arise throughout the European judicial area due to different choice-of-law rules in the Member States.⁵⁸ In applying their own conflict rules to ascertain the applicable law,⁵⁹ some Member States' courts refer to the *lex fori*, due to the procedural purpose of the jurisdiction clause,⁶⁰ while others apply the *lex causae*.⁶¹ Therefore, there is an inherent risk that the agreement's substantive validity may be assessed differently in different Member States.⁶² In turn, this raises the prospect of forum shopping, as a party may seise a particular court to have the jurisdiction agreement declared invalid.⁶³ In summary, the Brussels regime has been routinely criticised for not suitably protecting 'the sanctity of jurisdiction agreements'.⁶⁴ The Recast Regulation intends to address this criticism and attention now turns to the relevant amendments.

3. The Specific Legislative Reforms

The Recast supports the enforcement of choice-of-court agreements by substituting the current priority mechanism for an approach that grants precedence to the forum designated in the parties' agreement.⁶⁵ The priority rule is reversed by the cumulative effect of three strategic changes.⁶⁶ Firstly, Article 31(2) provides:

⁵⁴ Article 23(1) of the Regulation specifies the formal requirements of a choice-of court agreement and stipulates that the agreement must be: (a) in writing; or (b) in a form established in practice between the parties; or (c) in a form that accords with international trade usage.

⁵⁵ Case 24/76 *Salotti v RUWA* [1976] ECR 1831[7]. See also *Deutsche Bank AG v Asia Pacific Broadband Wireless Communications Inc* [2008] EWCA Civ 1091 [23] (Longmore LJ); Dickinson, 'Surveying the Proposed Regulation' (n 53) 285.

⁵⁶ Sancho Villa (n 33) 402.

⁵⁷ Adrian Briggs et al, *Dicey Morris & Collins on the Conflict of Laws* (14th edn, Sweet & Maxwell 2006) 12-108 cited in L Merrett, 'Article 23 of the Brussels I Regulation: a comprehensive code for jurisdiction agreements?' (2009) 58 ICLQ 545, 549.

⁵⁸ COM (2009) 174 final (n 4) 5.

⁵⁹ Steinle and Vasiliades (n 3) 578.

⁶⁰ Case C-269/95 *Benincasa v Dentalkit* [1997] ECR I-3767 [25].

⁶¹ COM (2009) 174 final (n 4) 5.

⁶² Sancho Villa (n 33) 403.

⁶³ Steinle and Vasiliades (n 3) 578.

⁶⁴ F Horn, 'Review of the Commission Proposal for a Recast of the Brussels I Regulation (COM [2010] 0748)' (*European Parliament, Directorate-General for Internal Policies, Legal Affairs*, 2011) <www.europarl.europa.eu/committees/de/studiesdownload.html?languageDocument=EN&file=49437> accessed 10 November 2012; Harris (n 18) 368-69; Sancho Villa (n 33) 402; Hartley, 'The Systematic Dismantling of the Common Law of Conflict of Laws' (n 41) 818.

⁶⁵ The Recast Regulation, Recital 22.

⁶⁶ Dickinson, 'Surveying the Proposed Regulation' (n 53) 290.

(...) where a court of a Member State on which an agreement referred to in Article 25 confers exclusive jurisdiction is seised, any court of another Member State shall stay the proceedings until such time as the court seised on the basis of the agreement declares that it has no jurisdiction under the agreement.

Secondly, the first-in-time rule is now located under Article 29 and explicitly states that this provision is 'without prejudice' to Article 31(2). Finally, Recital 22 emphasises that the parties' designated court is granted priority to establish jurisdiction regardless of whether it is first or second seised.⁶⁷

Another noteworthy change to the European jurisdictional area provided by the Recast Regulation relates to the rules on international *lis pendens*.⁶⁸ The Recast introduces two new provisions which grant the Member States' courts a discretion to stay proceedings in circumstances where the court of a non-Member State has already been seised: whereas Article 33 of the Recast operates where the proceedings involve the same cause of action and the same parties, Article 34 concerns conflicting proceedings involving related actions.⁶⁹ These novel provisions can also be considered exceptions to 'the otherwise strict operation' of the court-first-seised mechanism for allocating jurisdiction.⁷⁰

There are, however, important limitations to the application of the international *lis pendens* rules, namely: the conflicting proceedings must have been initiated first in the third State;⁷¹ the court of the third State is expected to render a judgment capable of recognition and enforcement in that Member State;⁷² and the Member State's court must be satisfied that staying the matter is necessary for the 'proper administration of justice'.⁷³ On this last limitation, Recital 24 of the Recast elaborates on the features that should be assessed by the court in exercising its discretionary power to suspend proceedings. Interestingly, the assessment has characteristics of the common law's *forum non conveniens* technique for declining jurisdiction in that the court is advised to take into account all of the circumstances of the case.⁷⁴ Such circumstances include, *inter alia*, the connections between the dispute, the parties and the third State, as well as the stage to which proceedings in the non-Member State have progressed.⁷⁵ Crawford and Carruthers observe that this exercise of judicial discretion is '(...) an unusual notion in a civilian inspired instrument'.⁷⁶ It is important to note, nonetheless, that the flexible mechanism provided by the Recast in relation to actions pending before the courts of third States

⁶⁷ A Layton, 'Comments by Alexander Layton QC on the proposed recast of the Brussels I Regulation' (European Parliament, Directorate-General for Internal Policies, Legal Affairs, 2011) <www.europarl.europa.eu/committees/en/juri/studiesdownload.html?languageDocument=EN&file=49443> accessed 25 November 2012.

⁶⁸ European Council (n 12) 1.

⁶⁹ Crawford and Carruthers (n 13) 92.

⁷⁰ *ibid.*

⁷¹ The Recast Regulation, Articles 33(1) and 34(1).

⁷² The Recast Regulation, Articles 33(1)(a) and 34(1)(b).

⁷³ The Recast Regulation, Articles 33(1)(b) and 34(1)(c).

⁷⁴ Crawford and Carruthers (n 13) 93.

⁷⁵ The Recast Regulation, Recital 24.

⁷⁶ Crawford and Carruthers (n 13) 93.

does not extend to the situation in *Owusu* because in that case the litigation in the Jamaican forum had not commenced.⁷⁷

Additionally, the legislators have subjected Article 23 of the Regulation, the current provision governing the prorogation of jurisdiction, to two important amendments. The first is that Article 25, the equivalent provision in the Recast, will operate irrespective of the parties' domicile.⁷⁸ The Recast has therefore effectively widened the scope of jurisdiction agreements captured by the instrument by removing the requirement in the existing Article 23 that at least one of the parties to the agreement is domiciled in a Member State.⁷⁹ The second amendment concerns a harmonised conflict of law rule that has been introduced to govern the substantive validity of the jurisdiction agreement.⁸⁰ The applicable law prescribed by the Recast's Article 25(1) is the law of the chosen forum.⁸¹ Specifically, this provision states that the chosen court shall have jurisdiction '(...) unless the agreement is null and void as to its substantive validity under the law of that Member State.'

The Commission considers that two of the main changes proposed by the Recast to the present jurisdictional scheme, the new choice-of-law rule and the revised priority ruling, '(...) reflect the solutions established in the HCCA and thereby [should] facilitate a possible conclusion to the Convention by the EU.'⁸² However, before discussing the Recast's alignment with the HCCA, it is important to analyse the impact of these reforms from both a regional and an international perspective.

4. Analysis of the Reforms

A. The Recast's *Lis Pendens* Rules

(i) The Revised Operation of the First-in-time Rule within Europe

The reversal of the priority ruling authorised by Article 31(2) of the Recast is primarily underscored by strong policy considerations, namely, to increase the effectiveness of jurisdiction agreements; discourage tactical proceedings;⁸³ and to promote the '(...) sound development of international commercial relations.'⁸⁴ Likewise, the theory supporting the revised priority rule is also sound. For instance, the drafting of the Recast's Article 31(2) is essentially based on Article 23(3) of the

⁷⁷ *ibid* 92. See also the Recast Regulation, Recital 23.

⁷⁸ Crawford and Carruthers (n 13) 92.

⁷⁹ S Garvey, 'Reform of the Brussels Regulation: are we nearly there yet?' (Allen & Overy LLP, 26 April 2013) <<http://www.allenoverly.com/publications/en-gb/Pages/Reform-of-the-Brussels-Regulation-are-we-nearly-there-yet.aspx>> accessed 31 May 2013.

⁸⁰ COM (2010) 748 final (n 10) 9.

⁸¹ Recital 20 of the Recast Regulation confirms that the applicable law which will be applied to determine the essential validity of the impugned jurisdiction agreement shall include both the law of the designated court and the conflict-of-law rules of that court's Member State.

⁸² COM (2010) 748 final (n 10) 9.

⁸³ The Recast Regulation, Recital 22. See also Fentiman (n 16) 247.

⁸⁴ AG Leger in Case C-116/02 *Erich Gasser GmbH v Misat Srl* 9 September 2003 ECR I-14693 [71].

Regulation which presently provides that when non-EU domiciliaries nominate a Member State court, all other courts shall have no jurisdiction over the dispute until the nominated court declines jurisdiction.⁸⁵ Article 23(3) clearly bestows '(...) great respect for the party autonomy of non-EU domiciliaries' that choose an EU court to resolve their dispute.⁸⁶ Beaumont avers that it is 'entirely logical' that the same respect should be extended to the autonomous choices of EU domiciliaries.⁸⁷

Yet reversing the supremacy attributed to the first-in-time rule is not without certain drawbacks. In particular, proceedings could be brought on the artifice of 'sham jurisdictional agreements'.⁸⁸ That is, by alleging the existence of a choice-of-court agreement and starting proceedings in the 'court supposedly designated', judgment will be delayed until jurisdiction is declined.⁸⁹ Horn argues that granting an '(...) unequivocal precedence for the allegedly-chosen court could in theory lead to "improved" torpedo actions'.⁹⁰ Despite Horn's claim, the allegedly-chosen court should be entitled to jurisdictional primacy on the basis that, in the many instances where the parties have negotiated a valid jurisdiction agreement,⁹¹ the possibility of thwarting the effectiveness of the agreement by pre-emptively striking in another court is negated.⁹² Furthermore, awarding the chosen court with first say in establishing jurisdiction '(...) fosters certainty and foreseeability in international commercial relations'.⁹³ To counter the possible threat of 'reverse' torpedo actions, however, appropriate costs sanctions could be introduced to limit the attraction of unconscionably raising a jurisdictional challenge.⁹⁴

Moreover, in all European jurisdictional conflicts under the Recast - even if the validity of the jurisdiction agreement cannot be supported - a solution is provided that is '(...) neutral to the order in which proceedings are brought'.⁹⁵ Most importantly, the temporal neutrality of the new policy reduces the parties' incentive to rush to the courthouse.⁹⁶ Notwithstanding the enhanced treatment of jurisdiction agreements in favour of Member State courts, there is disquiet over the application of jurisdictional clauses which nominate the courts of a third state. These concerns are the focus of the next section.

⁸⁵ Article 23(3) of the Regulation has been removed and, as indicated above, its substance is now found in Article 31(2) of the Recast.

⁸⁶ P Beaumont and P McEleavy, *AE Anton Private International Law* (3rd edn, W Green 2011) 259.

⁸⁷ *ibid.*

⁸⁸ Sancho Villa (n 33) 404. See also AG Leger (n 84) [74].

⁸⁹ *ibid* (AG Leger) [74].

⁹⁰ Horn (n 64) 24.

⁹¹ The Impact Assessment provides that a significant majority of EU corporations (70% of all companies and 90% of the larger companies that engage more than 250 employees) involved in cross-border trade incorporate choice-of-court agreements in their international contracts. See Impact Assessment (n 5) 28.

⁹² Beaumont and McEleavy (n 86) 259.

⁹³ *Gre Con Dimter Inc v JR Normand Inc* [2005] 2 SCR 401 cited in C McLachlan, 'Lis Pendens in International Litigation' (2008) *Recueil des Cours* 189, 248.

⁹⁴ Dickinson, 'Surveying the Proposed Regulation' (n 53) 296.

⁹⁵ Dickinson, 'The Proposed "Brussels I Bis" Regulation' (n 15) 19.

⁹⁶ *ibid.*

(ii) The Problems Associated with Third State Exclusive Jurisdiction Agreements and the International *Lis Pendens* Rules

Whilst Article 25 of the Recast applies regardless of whether the parties are EU domiciliaries, the scope of this provision remains limited to a choice of Member State courts.⁹⁷ In short, if a non-EU forum is designated in the parties' jurisdiction agreement, then such an agreement is not regulated by the Recast.⁹⁸

On this note, it is acknowledged that in conflicts of jurisdiction involving third States, the new international *lis pendens* rules in the Recast Regulation will offer some assistance to the Member States' courts in determining the venue for litigation.⁹⁹ However, it is submitted that in the interests of legal certainty, the ambit of the new regime should have been extended to include choice-of-court agreements designating non-Member States. A hypothetical case scenario illustrates the need for such an inclusion. If, for example, a party elected to defy an exclusive jurisdiction agreement in favour of Hong Kong by initiating an action in the UK (perhaps on the basis of the defendant's domicile),¹⁰⁰ then the Recast does not expressly permit the seised UK court to suspend its proceedings.¹⁰¹ The 'concession to discretionary reasoning'¹⁰² provided for in Articles 33 and 34 of the Recast is invoked if the proceedings in Hong Kong are on foot but the new international *lis pendens* rules do not cover those circumstances where the non-Member State court is seised second.¹⁰³ In essence, in the absence of a priority ruling for choice-of-court agreements in favour of third states, the UK court is compelled to accept jurisdiction and therefore acquiesce to the breach of contract.¹⁰⁴

It is possible that other issues could arise with respect to the Recast's mechanism for dealing with conflicting proceedings involving third States. First, the requirement in Articles 33 and 34 for the proceedings to have commenced first in a non-EU State creates an incentive to rush to court and could motivate defensive litigation strategies.¹⁰⁵ Secondly, at present the EU does not provide harmonised rules on the recognition and enforcement of non-Member State judgments. Accordingly, the possibility remains under the Recast regime for torpedo actions to be instigated by the simple expedient of bringing an action in the court of a Member State that does not recognise a judgment rendered in a third State.¹⁰⁶

⁹⁷ Crawford and Carruthers (n 13) 92.

⁹⁸ *ibid.*

⁹⁹ Garvey (n 79) 5.

¹⁰⁰ This jurisdictional base is pursuant to the general rule in Article 4 of the Recast, formerly Article 2 of the Regulation.

¹⁰¹ Garvey (n 79) 5.

¹⁰² Crawford and Carruthers (n 13) 92.

¹⁰³ The Recast Regulation, Articles 33(1) and 34(1).

¹⁰⁴ Garvey (n 79) 5.

¹⁰⁵ *ibid* 7.

¹⁰⁶ J Harris, 'Conflicts of Jurisdiction with Third States' (Brussels I Regulation Recast Conference, London, February 2013).

B. The Choice of Law Rule

The Commission asserts that the harmonised choice-of-law rule should promote legal certainty by ensuring uniform outcomes irrespective of the court seised.¹⁰⁷ Some commentators have questioned this assertion. Layton and Dickinson, for instance, allege that this proposed amendment to Article 23 of the Regulation could seriously impact on the case law developed by the CJEU on matters of validity.¹⁰⁸ It is important to note, however, that Article 25 of the Recast expressly introduces the words 'as to substantive validity' to prevent 'recourse to the choice-of-law rule' being applied to matters concerning formal validity.¹⁰⁹ In other words, reference to the applicable law is strictly limited to determining the substantive validity of a choice-of-court agreement. Thus, the CJEU's case law on the formal requirements of the Regulation's Article 23 will be preserved.¹¹⁰

Generally, jurists and stakeholders have been supportive of the proposed uniform choice-of-law rule.¹¹¹ Steinle and Vasiliades claim that procedural and economic benefits should accrue if the court designated in the putative jurisdiction agreement verifies the substantive validity of the agreement in accordance with its own law.¹¹² Such benefits relate to savings on time and court costs and predominantly arise because the external expertise of foreign lawyers is not required to advise on the laws of a foreign legal system.¹¹³

The endorsement provided by the UK's Ministry of Justice also exemplifies the support extended to the substantive validity choice-of-law rule.¹¹⁴ After observing that considerable disparity existed between the laws of the Member States on the conditions determining the essential validity of jurisdiction agreements, the Ministry insisted that it was 'most unlikely' that a consensus could be reached on a uniform interpretation of matters such as capacity, fraud or duress.¹¹⁵ The UK maintained that a harmonised choice-of-law rule most appropriately addressed the uncertainty which pervades the determination of substantive validity under the Regulation.¹¹⁶

This analysis suggests that the revised priority ruling and the uniform choice-of-law rule should increase the robustness of jurisdiction agreements within the EU. In addition, by holding the parties to their jurisdictional bargains, the proposed

¹⁰⁷ COM (2010) 748 final (n 10) 9.

¹⁰⁸ Layton (n 67) 15. See also Dickinson, 'Surveying the Proposed Regulation' (n 53) 300.

¹⁰⁹ Beaumont and McEleavy (n 86) 262.

¹¹⁰ Dickinson, 'Surveying the Proposed Regulation' (n 53) 286.

¹¹¹ Sancho Villa (n 33) 411. See also Beaumont and McEleavy (n 86) 262; European Economic and Social Committee, 'Opinion on the Proposal for a Regulation of the European Parliament and of the Council on jurisdiction and the recognition and enforcement of judgments in civil and commercial matters' COM (2010) 748 final - 2010/0383 (COD) (2011/C 218/14) [23.7.2011] [3.2] & [4.3] c/f Dickinson, 'The Proposed "Brussels I Bis" Regulation' (n 15) 21.

¹¹² Steinle and Vasiliades (n 3) 582.

¹¹³ *ibid.*

¹¹⁴ UK Ministry of Justice, 'Review of the Brussels I Regulation (EC 44/2001) Comments from the UK' (*European Parliament*, 3 September 2009).

<www.ec.europa.eu/justice/news/consulting_public/0002/contributions/ms_governments/united_kingdom_en.pdf> accessed 25 November 2012.

¹¹⁵ *ibid* [20].

¹¹⁶ *ibid.*

changes enhance the legal protection of European citizens.¹¹⁷ Accordingly, the Recast's two major amendments are, at the very least, meaningful. It is also necessary, however, to assess these reforms in light of the EU's intention to ratify the HCCA and the article will now address that requirement.

5. Alignment with the HCCA

The HCCA is an international treaty that applies to a specific group of transnational commercial disputes, namely, civil and commercial matters in which the parties have contracted for an exclusive choice-of-court agreement.¹¹⁸ Primarily, the Convention is premised on the assumption that such agreements have overriding value. The basic requirements on the courts of Contracting States are twofold. While the first is to validate party autonomy by upholding choice-of-court clauses, the second is to recognise and enforce the resulting judgments which emanate from the designated court.¹¹⁹ Most significantly, the HCCA is structured to make litigation outcomes more predictable which, in turn, promotes international trade and investment by reducing the risk associated with cross-border commercial transactions.¹²⁰

Put simply, the Convention's framework of rules is designed to ensure that a valid choice-of-court agreement confers exclusive jurisdiction to the parties' chosen court.¹²¹ Three core principles underpin this objective. First, Article 5(1) obligates the chosen court to assume jurisdiction.¹²² Critically, Article 5(2) reinforces the effectiveness of the parties' agreement by precluding the chosen court from resorting to the common law doctrine of *forum non conveniens* or applying the first-in-time rule.¹²³ Secondly, Article 6 holds that a seised court in a contracting state, which is not the chosen court, must refrain from exercising jurisdiction.¹²⁴ The final principle, one that Kruger considers 'goes to the very heart of the Convention',¹²⁵ is provided by Article 8(1) which specifies that the courts in other Contracting States must recognise and enforce a judgment delivered by the designated court.

¹¹⁷AG Leger (n 84) [3].

¹¹⁸ M Keyes, 'Jurisdiction Under the Hague Choice of Courts Convention: Its Likely Impact on Australian Practice' (2009) JPIL 181, 213.

¹¹⁹ L Teitz, 'The Hague Choice of Court Convention: Validating Part Autonomy and providing an Alternative to Arbitration' (2005) Am. J. Comp. L. 543, 547.

¹²⁰ *ibid* 557.

¹²¹ Sancho Villa (n 33) 405.

¹²² T Kruger, 'The 20th Session of the Hague Conference: a new Choice-of-Court Convention and the issue of EC Membership' (2006) 55 ICLQ 447, 451.

¹²³ T Hartley and M Dogauchi, *Explanatory Report on the Convention of 30 June 2005 on Choice of Court Agreements* (2007) Hague Conference on Private International Law [132].

<<http://www.hcch.net/upload/expl37e.pdf>> accessed 16 November 2012. It should be noted that although there is no room for the first-in-time rule in the HCCA, Article 19 does permit limited recourse to judicial discretion on the basis that there is no connection between that State and the parties or the dispute.

¹²⁴ Kruger (n 122) 451.

¹²⁵ *ibid* 452.

Hartley observes, however, that in situations of 'an exceptional nature',¹²⁶ the Convention authorises exemptions to the obligations detailed in Articles 6 and 8(1).¹²⁷ A common exception is that based on the validity of the jurisdiction agreement under the applicable law of the chosen court.¹²⁸ Hence, the three courts entitled to examine the validity of the jurisdiction agreement, namely, the designated court,¹²⁹ the non-chosen court¹³⁰ and the enforcing court,¹³¹ are compelled to review the matter in accordance with the same law - the law of the chosen court.¹³² Consequently, the harmonised applicable law is of paramount importance to the Convention because it minimises 'the possibility of irreconcilable judgments'.¹³³

Prima facie, this brief outline suggests that the Recast complies with the main structural provisions of the Convention.¹³⁴ The substance of the Recast's Article 31(2) accords with the provision in Article 5(2) of the Convention that the parties' chosen court is not obliged to stay proceedings if it was not the first court seised.¹³⁵ The Recast also adopts the uniform choice-of-law rule, the law of the chosen court, for assessing the substantive validity of the jurisdiction agreement.¹³⁶

The Recast has also been aligned with the Convention in relation to the scope of its application.¹³⁷ As previously indicated, Article 25(1) of the Recast applies 'regardless of domicile'. The test for application has therefore been reduced from the need to satisfy two conditions - domicile and the nominated court - to merely the requirement that an EU court is nominated.¹³⁸ This amendment in the Recast will 'ease the application of Article 26(6) of the Convention' in ascertaining whether the Regulation or the Convention applies to a particular dispute because the residence of the parties will solely determine which text applies.¹³⁹

The respective texts of the two instruments, however, are not identical. In fact, the obligations on the non-chosen courts are 'radically' different.¹⁴⁰ Pursuant to Article 31(2), the Recast strictly mandates that the non-chosen court must stay proceedings until the chosen court has established jurisdiction. Article 6(a) of the Convention, however, permits a non-chosen court to determine the validity and

¹²⁶ T Hartley, 'The Hague Choice-of-Court Convention' (2006) 31 EL Rev 414, 415.

¹²⁷ The five exceptions to Article 6 of the Convention are set out in Articles 6(a) - (e) while the grounds of refusal to the recognition or enforcement of judgments are established in Articles 9(a) - (g).

¹²⁸ The Convention, Articles 6(a) and 9(a) respectively. See also Hartley, 'The Hague Choice-of-Court Convention' (n 126) 416.

¹²⁹ The Convention, Article 5(1).

¹³⁰ The Convention, Article 6(a).

¹³¹ The Convention, Article 9(a).

¹³² Sancho Villa (n 33) 410.

¹³³ *ibid.*

¹³⁴ Sancho Villa (n 33) 417.

¹³⁵ *ibid.*

¹³⁶ It is important to note, however, that the two instruments differ in their approach to interpreting 'validity'. The role of formal validity is more important in the Regulation as consent in the Convention is not implied from the formality requirements. Instead, under the Convention, the existence of consent is governed by the law of the forum. See Hartley and Dogauchi (n 123) [94] - [96].

¹³⁷ Sancho Villa (n 33) 410.

¹³⁸ *ibid.*

¹³⁹ *ibid.*

¹⁴⁰ Beaumont and McEleavy (n 86) 259.

effect of a jurisdiction agreement.¹⁴¹ Further, the Convention's rules do not ban proceedings from commencing in the parties' chosen court even though '(...) a non-chosen court that was first seised of the dispute may still be hearing the case'.¹⁴² On this point, the Recast Regulation offers greater respect for party autonomy than the Convention.¹⁴³

The Recast Regulation, though, remains ambivalent on whether a Member State court is obliged to stay proceedings where an exclusive jurisdiction agreement designates a non-EU court.¹⁴⁴ This issue, intimated in *Owusu* but left open by the CJEU, remains uncertain. Recital 24 of the Recast attempts to shed some light on the issue and advises the Member States' courts that when exercising their discretion in conflicting proceedings involving third States, the assessment of the circumstances of the case may include '(...) consideration of the question whether the court of the third State has exclusive jurisdiction in the particular case'. In other words, a third State exclusive jurisdiction clause is merely one of the factors that a Member State court may consider in determining whether to stay proceedings. In contrast, the Convention respects exclusive choice-of-court agreements that are drafted in favour of any state.¹⁴⁵ In agreement with Dickinson, it is suggested that the EU's legislators should draft an amendment to the Recast based on Article 6 of the HCCA which obliges the non-chosen court to suspend or dismiss proceedings.¹⁴⁶

While the texts of the Recast and the Convention are not totally aligned, absolute alignment is not necessary nor, as Beaumont maintains, is it desirable.¹⁴⁷ Rather, the object of recasting the Regulation was to provide a coherent application of both instruments.¹⁴⁸ This objective has, in principle, been met and accordingly, the introduction of the Recast should not preclude the EU from ratifying the Convention.¹⁴⁹

6. Conclusion

The Recast improves the enforcement of exclusive jurisdiction agreements within the European legal framework by proposing two key amendments. Article 31(2) reverses the priority ruling which currently favours the first seised court while the substantive validity of jurisdiction agreements will be determined by a harmonised choice-of-law rule: the law of the chosen court. It is suggested that these reforms should achieve the primary objectives of the Recast, namely, to enhance the effectiveness of choice-of-court agreements and to ensure that such agreements are treated consistently in the Regulation and the HCCA.

¹⁴¹ The Convention, Article 6(a).

¹⁴² Beaumont and McEleavy (n 86) 259.

¹⁴³ *ibid* 260.

¹⁴⁴ Dickinson, 'Surveying the Proposed Regulation' (n 53) 302.

¹⁴⁵ *ibid*.

¹⁴⁶ *ibid*.

¹⁴⁷ Beaumont and McEleavy (n 86) 261.

¹⁴⁸ *ibid*.

¹⁴⁹ Sancho Villa (n 33) 417.

The Recast, however, could have contributed more to increase the overall effectiveness of choice-of-court agreements by addressing the *Owusu*-type situation where the designated court is a non-EU state. The EU legislators took steps towards improving the treatment of jurisdiction clauses in favour of third States with the introduction of the international *lis pendens* rules in Articles 33 and 34 of the Recast. Nevertheless, there are considerable limitations attached to the application of these provisions and a sense of uncertainty continues to be associated with conflicting proceedings involving non-Member States' courts. Notably, the potential for tactical litigation remains under the latest instalment of the Brussels regime. In light of these concerns, the inclusion of an express provision in the Recast Regulation granting precedence to the court of a non-Member State designated in an exclusive jurisdiction agreement should be advocated to augment legal certainty in the European judicial area.

Finally, the civilian and common law traditions adopt fundamentally different approaches to resolving conflicts of jurisdictions.¹⁵⁰ The formalism at the core of the *lis pendens* doctrine promotes procedural efficiency by negating concurrent proceedings.¹⁵¹ The common law, on the other hand, considers that the rules of jurisdiction are designed to protect the interests of the parties.¹⁵² Crucially, civil law concepts have largely influenced the CJEU and it is open to suggestion that the decisions in *Gasser* and *Owusu* were tarnished by the court's fidelity to structural and procedural certainty.¹⁵³ Accordingly, the proposed jurisdictional agreement reforms, which acknowledge the significance of party autonomy and predictability in international commercial transactions,¹⁵⁴ qualify as necessary and meaningful. It is suggested, moreover, that in striking a balance between procedure and justice, the Recast has, correctly, shifted the equilibrium point towards the common law's position.

¹⁵⁰ *Airbus Industrie GIE v Patel* [1999] 1 AC 119, 132-33. See also Hartley, 'The Systematic Dismantling of the Common Law of Conflict of Laws' (n 41) 814; Gardella and Radicati di Brozolo (n 22) 611-14.

¹⁵¹ Hartley, *International Commercial Litigation* (n 40) 241.

¹⁵² Harris 'Understanding the Europeanisation of Private International Law' (n 18) 372.

¹⁵³ Hartley, 'The Systematic Dismantling of the Common Law of Conflict of Laws' (n 41) 820-21, 826-27.

¹⁵⁴ *ibid* 817-18.

Gender Mainstreaming & 'Equality Proofing' in British Law-Making: A Comment on the Impact of the Equality Act 2010

ILONA C.M. CAIRNS*

Abstract

The central purpose of this article is to demonstrate that 'gender mainstreaming' and 'equality proofing' procedures in British law-making are in a state of flux. It reflects on two recent developments that are likely to have a significant impact on how gender equality concerns are taken into account in law-making: the introduction of the single public sector equality duty in the Equality Act 2010 and the announcement in late 2012 by David Cameron that public authorities are no longer required to carry out 'equality impact assessments'. Although both of these developments threaten to undermine the fundamental purpose of gender mainstreaming, a widely-endorsed equality strategy that requires all law and policy to be evaluated through a gendered lens, they also send conflicting messages to public authorities. Whereas the new public sector equality duty requires law-makers to be more aware of diversity and to take more equality concerns into account in law and policy making, the axing of equality impact assessments reflects the view that such practices are overly bureaucratic and a waste of valuable resources. This paradox exacerbates the lack of clarity that currently defines 'equality proofing' in British law-making and raises new and serious questions about the responsiveness of future laws to the needs and interests of diverse social groups.

1. Introduction

In April 2011, the Equality Act 2010 came into force across Great Britain,¹ replacing specific public sector equality duties relating to gender, disability and race with a single public sector equality duty.² This new 'general' duty requires *inter alia* that public authorities 'have due regard' to the elimination of discrimination, harassment and victimisation³ directed at individuals with a 'protected characteristic'⁴ and further that such authorities 'advance equality of opportunity' and 'foster good relations' '(...) between persons who share a relevant protected characteristic and

* PhD Candidate, School of Law, University of Aberdeen.

¹ With limited exceptions, the Act does not apply to Northern Ireland. See Equality Act 2010, s 217.

² <<http://www.scotland.gov.uk/Topics/People/Equality/18500/GenderEqualityIssues>> accessed 23 May 2013; Equality Act 2010, s 149.

³ Equality Act 2010, s 149 (1) (a).

⁴ Section 149 (7) of the Equality Act 2010 lists the 'protected characteristics'. These are age, disability, gender reassignment, pregnancy and maternity, race, religion or belief, sex and sexual orientation.

persons who do not share it.⁵ Approximately eighteen months after the Equality Act came into force, David Cameron announced that 'equality impact assessments' (EIAs), a tool used to evaluate the impact of policy and legislation on protected groups, were to be axed.⁶ This announcement appeared to send the message that the evaluation of legislation and policy against a standard of equality is unnecessary 'red tape'⁷ rather than a necessary and prudent measure to ensure that British laws do not have a disparate impact on society's most vulnerable groups. With respect to gender equality, Cameron's announcement is in direct conflict with the concept of gender mainstreaming ('GM') which demands that '(...) *any* planned action, including legislation, policies or programmes, in *any* area and at *all* levels'⁸ be evaluated for their potential gendered impact.

This article analyses the above two developments in light of the UK's ostensible commitment to gender mainstreaming, paying specific attention to how and why gender equality concerns are currently taken into account in the law-making process. In order to comment upon the possible consequences of removing EIAs (themselves a tool of gender mainstreaming), it is necessary to first examine the arguments justifying their existence. After providing an overview of the concept of gender mainstreaming, part two of this article outlines the arguments in favour of gender mainstreaming in general and gender equality proofing specifically. The following part then draws attention to concerns that have been expressed with respect to the implementation of gender mainstreaming, questioning whether the axing of EIAs will in fact make a material difference in terms of achieving gender equality.

The final part of this article reflects on the meaning and potential impact of the new public sector equality duty. It will be argued that although the new law means that gender is no longer singled out as requiring specific attention in policy formation and legislative drafting, the new duty signals a movement towards 'diversity' or 'equality' mainstreaming that has been recently popularised and promoted in both mainstreaming and feminist literature. In short, the disappearance of the gender equality duty from Britain's equality law framework is not necessarily problematic from the perspective of those most concerned with the advancement of gender equality. The adoption of a single equality duty does, however, raise a wealth of new concerns that, coupled with the abolition of EIAs, cloak the future of equality proofing in British law-making in a cloud of uncertainty.

⁵ Equality Act 2010, s 149 (1) (b) & (c).

⁶ H Mulholland, 'David Cameron axes equality assessments in war on 'red tape' *The Guardian* (London 19 Nov 2012) <<http://www.theguardian.com/politics/2012/nov/19/cameron-axe-equality-assessments>> accessed 23 May 2013.

⁷ *ibid.*

⁸ This is part of the commonly-referred to definition of gender mainstreaming, coined by the UN Economic and Social Council in 1997. See United Nations, *Report of the Economic and Social Council for 1997* (New York: General Assembly, A/53/3) available at <<http://www.un.org/documents/ga/docs/52/plenary/a52-3.htm>> accessed 23 May 2013.

2. The Argument for Mainstreaming Gender Equality in British Law & Policy-Making

A. Gender Mainstreaming in England, Wales & Scotland

To many, the notion that all law and policy should be evaluated for its gendered impact might sound at once controversial, overtly feminist, radical and impractical. However, since gender mainstreaming was formally adopted in 1995 at the UN World Conference on Women in Beijing,⁹ and by the EU in the Treaty of Amsterdam,¹⁰ this idea has been officially embraced and promoted on both an international and domestic level. The Council of Europe have offered the following justificatory rationale for the universal adoption of GM:¹¹

Gender mainstreaming is essential for a properly functioning democracy. It puts people at the heart of policy making; leads to better informed policy-making and therefore enhanced government; makes full use of all human resources and acknowledges the shared responsibilities of women and men in all spheres of social ordering; makes gender visible at all levels of society; and takes account of diversity between women and men and between women and women, men and men.

Significantly, both the Scottish and UK Parliaments have endorsed GM as a 'gender equality strategy',¹² with the Scottish Executive describing equality mainstreaming as '(...) the systematic integration of an equality perspective into the everyday work of government, involving policy makers across all government departments, as well as equality specialists and external partners.'¹³ It is important to note here that although the Scotland Act reserves the power to legislate on equal opportunities to the UK Parliament, the Scottish Executive still has the power to 'encourage equal opportunities' and impose specific duties on Scottish public authorities.¹⁴ There are therefore some key differences in the mainstreaming duties

⁹ O Hankivsky, 'Gender vs. Diversity Mainstreaming: A Preliminary Examination of the Role and Transformative Potential of Feminist Theory' (2005) 38:4 Canadian Journal of Political Science 977, 978.

¹⁰ The Treaty of Amsterdam 1997, Articles 2 & 3. Kantola points out that the Council of Europe's 2008 definition of gender mainstreaming has been influential in the European context. See J Kantola, *Gender and the European Union* (Palgrave 2010) 127. The Council of Europe defines GM as '(...) the (re)organization, improvement, development and evaluation of policy processes, so that a gender equality perspective is incorporated in all policies and all levels and at all stages, by the actors normally involved in policy making.' Cited in Kantola, 127.

¹¹ Council of Europe, cited in C Chinkin, *Gender Mainstreaming in Legal & Constitutional Affairs: A Reference Manual for Governments and Other Stakeholders* (Commonwealth Secretariat 2001) 14.

¹² Equal opportunities and positive action are generally recognized as the two other 'equality' strategies. See R Bendl & A Schmit, 'Gender Mainstreaming: An Assessment of Its Conceptual Value for Gender Equality' (2013) 20:4 Gender, Work & Organisation 364, 364.

¹³ <<http://www.scotland.gov.uk/Topics/People/Equality/18507/executivemainstreaming>> accessed Aug 2 2013.

¹⁴ <<http://www.scotland.gov.uk/Topics/People/Equality/18507/govtresponsibility>> accessed Aug 2 2013. The Equality Act makes provision for specific duties to be laid down by Scottish Ministers through the adoption of regulations. The Equality Act 2010 (Specific Duties) (Scotland) Regulations

imposed on public bodies in Scotland and the rest of the UK. Indeed, some have suggested that '(...) the commitment to mainstreaming (...) appears to be a stronger force for change in the newly devolved administrations of Scotland and Wales' due to the heavy involvement of women's groups in the devolution process. As McKay et al explain:¹⁵

The absence of a stated political commitment to gender in election manifestos suggests that the promotion of gender balance and gender mainstreaming in Scotland can be directly attributed to the lobbying and participants of women's groups throughout the process towards devolution.

Despite some differences in how gender mainstreaming is conceived and implemented, however, the fact that the fundamental equality law framework is governed by Westminster means that legislative change also applies in the devolved administrations of the UK.

B. The Logic Underlying the Gender Equality Duty

The foundational idea of gender mainstreaming is that state policies and laws affect women and men differently¹⁶ and thus have the potential to reproduce and perpetuate existing patterns of gender inequality. The widespread and systematic evaluation of law and policy (both international and domestic) for its potential differential impact on men and women is the core aim of GM. This aspiration is strikingly wide, but also 'radical'¹⁷ for, as Zalewski has noted:¹⁸

(...) the generic demand [of GM] is nothing short of wholesale transformation of the institutions and processes of government in regard to gender, with the intention that this impacts on, works with, and changes the wider society and polity.

GM has been described '(...) as the most "modern" approach to gender equality'¹⁹ insofar as it moves beyond a focus on individual rights²⁰ towards an approach that challenges '(...) those systems, processes and norms that generate inequalities'.²¹ Gender mainstreaming might also be understood as 'modern' insofar as it attempts to shift the conceptual focus away from 'women's' interests towards more 'generic'

came into force in May 2012. The regulations are available online:

<www.legislation.gov.uk/sdsi/2012/9780111016718/contents> accessed Aug 2 2013.

¹⁵ A McKay, R Fitzgerald, A O'Hagan & M. Gillespie, 'Scotland: Using Political change to advance gender concerns' in D Budlender & G Hewitt (eds), 'Gender Budgets Make More Cents: Country Studies and Good Practice' (Commonwealth Secretariat 2002).

¹⁶ Hankivsky (n 9) 977.

¹⁷ M Zalewski, 'I don't even know what gender is': a discussion of the connections between gender, gender mainstreaming and feminist theory' (2010) 36 *Review of International Studies* 3, 7.

¹⁸ *ibid.*

¹⁹ M Daly, 'Gender Mainstreaming in Theory and Practice' (2005) 12:3 *Social Politics: International Studies in Gender, State & Society* 433, 433.

²⁰ Bendl & Schmit (n 12) 364.

²¹ *ibid.*

gender issues. Otherwise put, GM is not 'supposed' to be a strategy only for women, despite the fact that the concept has strong connections with the feminist movement.²² Rather, GM was conceived as a strategy to ensure that the needs and interests of women *and* men are evaluated and responded to in law and policy making. Indeed, some have argued that the focus on 'gender' rather than 'women' is a key advantage of GM;²³ that the fact that mainstreaming is not explicitly 'feminist' helps to '(...) win broader audiences for gender issues'.²⁴

Other commentators have highlighted that, in reality, gender mainstreaming is often treated as if it is really about women or, to put it another way, '(...) that gender mainstreaming may not necessarily be gender-focused at all.'²⁵ In the UK context, this is exemplified by the fact that the UK's first response to the advent of GM was for the Ministers of Women to set up a Women's Unit within the Department of Work and Pensions.²⁶ This is likely explained by the fact that women are still framed as the 'problem holders' in gender inequality discourse.²⁷ Nevertheless, it is the introduction of the Gender Equality Duty (the 'GED') in 2006 that still stands out as the UK government's most obvious and concerted effort to mainstream gender.²⁸ Indeed, at the time of its introduction, the GED was heralded as the most significant legislative development in the area of gender equality since the Sex Discrimination Act.²⁹ Acknowledging the failure of an 'individual rights' approach, the GED placed the onus on public authorities to integrate gender considerations into their everyday work, requiring them to publish gender equality schemes and, importantly, '[t]o assess the impact of its current and proposed policies and practices on gender equality, and to have due regard to the results of those impact assessments'.³⁰ The duty applied '(...) to policy-making, service provision, employment matters, and in relation to enforcement or any statutory discretion and decision making.'³¹ It is therefore clear that the full and proper implementation of the GED demanded that gender equality concerns be paid 'due regard' in the formulation and drafting of law and policy. A key question, of course, is how to measure the gender equality impact of proposed legal and policy changes. What

²² As Hankivsky has noted, '[f]eminist theories about engagement with the state and normative arguments regarding women's oppression, subordination and inequality constitute the foundation on which GM is constructed'. Hankivsky (n 9) 983.

²³ E Woodward, 'European Gender Mainstreaming: Promises and Pitfalls of Transformative Policy' cited in Hankivsky (n 9) 984.

²⁴ Hankivsky (n 9) 984.

²⁵ Daly (n 19) 441. See also Zalewski (n 17) 6, noting that gender mainstreaming has been accused of ignoring men and masculinity.

²⁶ J Veitch, 'Looking at gender mainstreaming in the UK government' (2006) 7:4 International Feminist Journal of Politics 600, 600. Harriet Harman and Joan Ruddock were the Ministers for Women in 1997.

²⁷ E Lombardo & P Meier, 'Gender Mainstreaming in the EU: Incorporating a Feminist Reading?' (2006) 13: 2 European Journal of Women's Studies 151, 157.

²⁸ The GED was introduced by the Equality Act 2006, which amended the Sex Discrimination Act 1975.

²⁹ See Equality & Human Rights Commission (Scotland), *Gender Equality Duty Code of Practice: Scotland* 2 available at <www.equalityhumanrights.com/.../scottish_ged_code_of_practice.doc> accessed Aug 4 2013. See also Zalewski (n 17) 17.

³⁰ Equality & Human Rights Commission (Scotland) (n 29) 6-7.

³¹ *ibid* 6.

exactly is involved in equality proofing and what are some of the limits and failures of mainstreaming gender in law-making?

3. Mainstreaming Gender: How it Works in Practice

The integration of gender equality considerations can occur at various stages in the law-making process and via a variety of mainstreaming 'tools', which include equality proofing procedures and gender-based analysis.³² Equality impact assessments form an important part of the wider process of equality proofing, however it is important to note that they were not legally required under the gender equality duty, nor are they currently required under the public sector equality duty.³³ On the other hand, they are carried out by public authorities as a matter of good practice and serve as *proof* that public authorities have fulfilled their legal duty to pay 'due regard' to the elimination of discrimination and the promotion of equality.³⁴ Their purpose is to '(...) expose the 'gap' between the assumptions on which policy has been based and the reality',³⁵ rather than to ensure blanket gender neutrality or '(...) that all decisions are "good for women"'.³⁶ The following passage, taken from a Scottish Executive report that compared equality proofing procedures in different jurisdictions, contains a useful summary of the types of issues that mainstreaming tools can expose in law and policy formation:³⁷

(...) flaws have been exposed in proposed policies and legislation. For example, gender impact assessments carried out at policy review stage in the Netherlands showed that a plan to restructure secondary school education which the designers believed to be 'gender neutral' would in fact reinforce gender segregation within the educational system. Similarly, when an analysis of a proposed reform of the electoral system was conducted, it showed that it would actually reduce the number of elected women politicians. In health

³² F Mackay & K Bilton, *Equality Proofing Procedures in Drafting Legislation: International Comparisons* (Scottish Executive Central Research Unit 2001). At 6, the report explains that these terms are often used interchangeably, however notes that it is more accurate to think of equality proofing '(...) as the formal framework within which impact assessment takes place.'

³³ This matter was discussed in the case of *R (Brown) v Secretary of State for Work and Pensions and another (Equality and Human Rights Commission intervening)* [2008] EWHC 3158 (Admin), cited in D Pyper, 'The public sector equality duty and equality impact assessments' (Briefing Paper, House of Commons Library, 21 March 2013) 7 available at <www.parliament.uk/briefing-papers/SN06591> accessed Aug 6 2013.

³⁴ *ibid.* At 17, Pyper notes that '(...) although the law does not require public authorities to carry out EIAs, the courts place significant weight on the existence of some form of documentary evidence of compliance with the PSED when determining judicial review cases.'

³⁵ Mackay & Bilton, *Equality Proofing Procedures in Drafting Legislation* (n 32) 5.

³⁶ *ibid.*

³⁷ *ibid.* The report also contains examples of the types of questions that may be asked in the process of assessing whether a particular law or policy has relevance to gender. One such example (outlined on pgs. 8-9) is the Dutch analytical tool SMART (Simple Method to Assess the Relevance of policies to gender) which is comprised of two questions: '1. Is the policy proposal directed at one or more target groups? Will it affect the daily life of part(s) of the population? 2. Are there differences between women and men in the field of the policy proposal (in respect of rights, resources, representation, values and norms related to gender?'

care, a strategy which had been proposed to improve services for those suffering from chronic illnesses had taken as its reference point the needs and lives of young men, whereas in fact, most of the chronically sick are older women.

The report also found that the effectiveness of GM was maximised if equality proofing and/or gender-based analysis were conducted at various stages, including policy formation, legislative drafting and scrutiny of the Bill.³⁸ At the same time, however, the report suggests that the likelihood of integration is increased if proofing takes place in the early stages of policy formation and is followed by '(...) an ongoing series of checks and interventions.'³⁹ It is perhaps unsurprising that the reality falls short of this ideal. Due to limited resources and time constraints, screening methods are commonly used to prioritise what 'needs' to be equality proofed. Although it does not appear that any 'official' screening methods exist in Britain, it seems likely that a degree of filtering occurred/occurs⁴⁰ through the requirement that public authorities pay 'due regard' to the promotion of equality and the potential for discrimination. The Scottish Code of Practice for the GED explained that 'due regard' comprises the elements of 'proportionality' and 'relevance', stating that '[i]n practice, this principle will mean public authorities should prioritise action to address the most significant gender inequalities within their remit.'⁴¹ The Equality Act 2010 does not remove the uncertainty surrounding what constitutes 'due regard' and concern has recently been expressed over the vague and non-specific nature of the term.⁴² Although some guidance can be found in case law,⁴³ to date there is no official code or guidance providing clear instructions to public authorities on this issue, meaning that those subject to the public sector duty, as well as the courts, have a degree of discretion as to what is prioritised.

While prioritisation is certainly understandable, there is a risk that the fundamental purpose of mainstreaming is undermined through screening or filtering processes. As MacKay and Bilton note, screening means that legislation or policy that appears 'gender neutral' or does not have *obvious* implications for gender equality will be screened out, despite the fact that it might have considerable gendered implications.⁴⁴ This goes against the very ethos of gender 'mainstreaming'. The all-embracing demands of GM might also mean that mainstreaming tools such as equality proofing are employed in a routine, essentially meaningless fashion, with

³⁸ Equality proofing can be conducted by a variety of public servants, including '(...) policy makers, equality experts within departments, specialist equality units, Bill teams including instructing officers and drafters, statutory equality agencies, parliamentary committees and external experts.' *ibid* 6.

³⁹ *ibid* 7.

⁴⁰ As noted above, the new legislation relating to the public sector duty also uses the term 'due regard.'

⁴¹ Equality & Human Rights Commission (Scotland) (n 29) 16.

⁴² In a recent lecture, Justice Sales said the following: 'The statute does not give us much information about that [what is 'due regard'], other than again in very general terms in section 149(3). The practical effect of the combination of a very wide range of application for the duty across all public functions and a very abstract formulation of what has to be done means that the burden of spelling out the practical content of the duty devolves to the courts.' Cited in Pyper (n 33) 7.

⁴³ *ibid*. See, e.g., *R (Brown) v Secretary of State for Work and Pensions and another* (n 33).

⁴⁴ Mackay & Bilton, *Equality Proofing Procedures in Drafting Legislation* (n 32) 8.

those responsible for proofing adopting a 'tick-box' mentality. Finally, it is worth bearing in mind that all UK legislation must, at a minimum, comply with anti-discrimination, equality and human rights legislation, a fact that might lead some law-makers to conclude that sufficient equality proofing has already occurred, or will occur in the future, thereby deterring further analysis. All of the above factors cast doubt over the true value of equality proofing.

Difficulties relating to the systematic implementation of equality proofing mirror the broader difficulties with GM that are well-reported in mainstreaming literature. In her analysis of an EU-funded research project (EQUAPOL) that examined how gender is being mainstreamed in eight European countries,⁴⁵ Daly reported enormous disparity and fragmentation in approaches to mainstreaming, concluding that the '(...) "common core" to gender mainstreaming in action across countries (...) lies in the tendency to apply the approach in a technocratic way and to be non-systemic in compass.'⁴⁶ Other mainstreaming commentators paint a similar picture with respect to difficulties with implementation. While Hankivsky claims that 'GM's promise (...) has not been realized in any jurisdiction or in any area of public policy',⁴⁷ Squires has endorsed the view that '(...) many countries and organizations adopt mainstreaming in name only'.⁴⁸ Recently, there have even been signs of a movement away from GM in the EU, leading some to declare a state of 'crisis'.⁴⁹ With this in mind, David Cameron's view that EIAs - which as noted above form part of the proofing procedure encouraged by mainstreaming - are nothing more than 'bureaucratic rubbish'⁵⁰ does not seem quite so extreme. As with mainstreaming generally, there is no available evidence to support the view that impact assessments *alone* will deliver practical improvement to the lives of those with protected characteristics. The research conclusions of MacKay and Bilton support this view:⁵¹

Whilst formal requirements to include statements about gender or other equality impact assessments in Memoranda for Cabinet are seen as desirable, they are not seen as sufficient to ensure effective integration of such considerations in legislative proposals. If these requirements are not backed up by mainstreaming systems, training, resources, good working relationships and political will they are seen as symbolic rather than resulting in changes outcomes.

⁴⁵ Daly (n 19) 434. The countries were Belgium, France, Greece, Ireland, Lithuania, Spain, Sweden and the United Kingdom.

⁴⁶ *ibid* 433.

⁴⁷ Hankivsky (n 9) 981.

⁴⁸ J Squires, *The New Politics of Gender Equality* (Palgrave 2007) 73. Here, Squires concludes: 'The inconsistent and unenthusiastic application of mainstreaming may be symptomatic of its derivation in institutions beyond the nation-state, meaning that the strategy lacks the necessary internal advocates, fails to mesh with existing policy priorities or institutional structures, or becomes interpreted in ways which resonate with national priorities, norms and structures, but which limit its envisaged potential.'

⁴⁹ Bendl & Schmit (n 12) 364.

⁵⁰ Mulholland (n 6).

⁵¹ MacKay & Bilton, *Equality Proofing Procedures in Drafting Legislation* (n 32) 41.

Nevertheless, it is suggested here that the outright axing of equality impact assessments is overly hasty, out of step with the approach taken in other EU and Commonwealth countries and sends the wrong message about the government's commitment to equality. After all, it seems clear that the difficulties with EIAs pertain to their implementation and effectiveness, as opposed to the reasoning and values underpinning their existence. The decision to outright abandon impact assessments, rather than to dedicate more energy and resources into creating a culture in which mainstreaming is valued and supported, has therefore been criticised. Former shadow equalities minister, Yvette Cooper, raised concerns about axing impact assessments at a time when '(...) women are being hit much harder than men with low-income working families and disabled families losing out badly',⁵² accusing the prime minister of removing '(...) any requirement for the public sector to even think about equality' and of having a 'personal blind spot on women.'⁵³ Even if there is not yet available evidence pointing to a direct correlation between impact assessments and measurable equality outcomes, it is argued here that EIAs (and equality proofing generally) are part and parcel of fair, transparent and responsible law-making procedures that define a democratic state. Even if a degree of bureaucracy is involved, this is preferable to placing the burden on already over-worked public servants to judge when equality concerns are relevant and worthy of further analysis.

Although lack of political will may go some way to explaining the lack of effectiveness with respect to equality proofing procedures in Britain, it is important not to lose sight of the fact that the failure of mainstreaming is a universal problem, even in countries such as Canada and Sweden where there is strong political support and clear proofing guidelines and procedures in everyday use.⁵⁴ The fact that there is still ambiguity as to exactly what GM entails has prompted deep reflection on the exact meaning of the terms 'gender', 'mainstreaming' and 'equality', as well as on the theoretical underpinnings of GM. Implementation issues have also triggered debate about how to 'reinvigorate' GM, with many arguing for a more diversity-focused concept of mainstreaming. As noted earlier in this article, the new public sector equality duty in the Equality Act 2010 more closely conforms to a diversity model of mainstreaming than the previous individual equality duties. The remainder of this article assesses the significance of this change, again using insights from mainstreaming literature.

4. Diversity Mainstreaming & The Equality Act 2010

A. The Shift Towards 'Diversity' or 'Equality' Mainstreaming

Although the former specific public sector equality duties relating to gender, race and disability were presented as powerful, revolutionary, tools at the time of their introduction, they were promptly criticised on grounds of inadequacy. From a

⁵² Mulholland (n 6).

⁵³ *ibid.*

⁵⁴ See generally MacKay & Bilton, *Equality Proofing Procedures in Drafting Legislation* (n 32).

conceptual standpoint, the principal critique was that these distinct duties reflected a narrow understanding of inequality, one that required individuals to identify with only one homogenous group in order to pursue a case of discrimination, even if those individuals had experienced discrimination on multiple levels.⁵⁵ For example, a black disabled woman would have had to *choose* the ground on which she wanted to pursue a claim, when in reality the discrimination might have been triggered by the intersection of her different disadvantages.⁵⁶ As Bagihole explains in some detail, UK statistics show that social inequality is experienced in complex ways. For example, while disabled men and women experience higher levels of unemployment compared to those who are not disabled, disabled women are less likely to be employed than disabled men.⁵⁷ The fact that Muslim women suffer high levels of economic disadvantage compared with both women and men in different religious groups⁵⁸ also serves as evidence of '(...) diverse and intersectional (in)equality between differentiations and communities, and polarization within each.'⁵⁹

Numerous mainstreaming commentators have therefore drawn attention to the difficulties with equality strategies that '(...) look at gender equality in isolation from other forms of equality',⁶⁰ calling for equality strategies that are responsive to diverse forms of inequality, both between and apart from men and women. Hankivsky has presented a strong case for reconceptualising GM to be more diversity-focused, arguing that '(...) GM invokes a liberal concept of an abstract woman'⁶¹ that '(...) tends to concentrate on differences between men and women, treating each gender as a unitary, one dimensional category of analysis'⁶² in a way that nourishes '(...) fairly crude distinctions between women and men.'⁶³ As with GM theory as a whole, these claims draw heavily on the work of some feminist writers who have taken opposition to the notion that all women share the same 'voice', needs, problems and experiences,⁶⁴ and have popularised the term 'intersectionality'.⁶⁵

Since the mid-2000s, therefore, calls to recognise and respond to plural, intersecting forms of inequality have gained in frequency and encouraged a widespread shift towards what is commonly termed 'diversity' or 'equality'

⁵⁵ B Bagilhole, 'Equal Opportunities and Diversity Policies' (2010) 27 Canadian Journal of Administrative Sciences 263, 263.

⁵⁶ *ibid.*

⁵⁷ *ibid* 265.

⁵⁸ *ibid.*

⁵⁹ *ibid.*

⁶⁰ J Squires, 'Is Mainstreaming Transformative? Theorizing Mainstreaming in the Context of Diversity and Deliberation' (2005) 12:3 Social Politics 366, 367.

⁶¹ Hankivsky (n 9) 986.

⁶² *ibid.*

⁶³ *ibid.*

⁶⁴ See, e.g., JG Greenberg, 'Introduction to Postmodern Legal Feminism' in MJ Frug, *Postmodern Legal Feminism* (Routledge 1992) x. As Greenburg notes here, such concerns about the meaning and scope of the term 'women' remind us that '(...) generalized perspectives are necessarily partial' and that defining '(...) one particular group of women as representing the "essence" of women does violence to and constrains the lives of women who differ.'

⁶⁵ The work of Crenshaw was groundbreaking with respect to intersectionality. See, e.g., K Crenshaw, 'Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Colour,' (1991) 43:6 Stan.L.Rev. 1241.

mainstreaming. As Bacchi and Eveline note, the language of diversity has been embraced by the EU, as well as by legal international organisations, including the United Nations and the World Bank.⁶⁶ The introduction of the unitary public sector equality duty in the Equality Act 2010 is the legal manifestation of this shift in Scotland and England and Wales. However it was the establishment of the Equality and Human Rights Commission (EHRC) three years prior to this Act that clearly marked the adoption of a 'single equality approach' in Britain.⁶⁷ While it is almost impossible to criticise the reasoning behind the extension of equality protection to other disadvantaged groups, praise for Britain's new single equality approach has not been universal. This article now highlights several unintended consequences that have the potential to undermine the transformative potential of the new equality approach embodied in the public sector equality duty.

B. The Public Sector Equality Duty: Too Much, Too Early?

In presenting her case for 'diversity' mainstreaming, Hankivsky makes the following claim:⁶⁸

(...) diversity mainstreaming allows for a more complex and dynamic understanding of equality and social justice, because the contours and compound effects of discrimination that women experience can be captured and the invisibility or marginalization of difference is no longer an option.

Hankivsky is thus of the view that 'diversity' mainstreaming creates an environment that is *more* responsive to *more* women's needs. There is a strong argument, however, that the opposite is true in practice; that the requirement that public authorities evaluate the equality impact of law and policy on numerous grounds may lead to gender equality 'losing out' in the 'milieu of diversity.'⁶⁹ Several mainstreaming commentators have voiced concern along these lines. In their analysis of the reform of European employment strategy, for example, Fagan, Grimshaw and Rubery argue that gender mainstreaming and gender equality objectives have been 'subordinated' following the removal of gender equality guidelines.⁷⁰ Although the new public sector equality duty does not remove the duty on public authorities to take gender issues into account,⁷¹ it is possible to argue that the new duty obscures gender considerations to a certain extent and removes them from the forefront of the mind of law and policy makers. This is not to argue that gender is somehow more important or more worthy of attention than other protected characteristics, but simply that the

⁶⁶ C Bacchi & J Eveline, 'Gender Mainstreaming or Diversity Mainstreaming? The Politics of "Doing"' (2009) 17:1 *Nordic Journal of Feminist and Gender Research* 2, 6.

⁶⁷ E McLaughlin, 'From Negative to Positive Equality Duties: The Development and Constitutionalisation of Equality Provisions in the UK' (2007) 6:1 *Social Policy & Society* 111, 111.

⁶⁸ Hankivsky (n 9) 994.

⁶⁹ Bagilhole (n 55) 265.

⁷⁰ C Fagan, D Grimshaw & J Rubery, 'The subordination of the gender equality objective: The National Reform Programmes and 'making work pay' policies' (2006) 37:6 *Industrial Relations Journal* 571.

⁷¹ Although interestingly the term 'gender' is replaced with 'sex' in the new Equality Act 2010.

generic equality analysis approach, encouraged by the Equality Act 2010, carries the risk of '[d]ilution and blandness'.⁷²

There are other potential downfalls associated with the UK's new approach to equality. Verloo has helpfully identified three basic difficulties relating to the EU's movement towards a multiple inequalities approach that are also relevant to immediate discussion: '(...) the assumed similarity of inequalities, the need for structural approaches and the political competition between inequalities'.⁷³ The first concern stems from the way that single equality instruments (e.g. the Equality Act 2010) list inequalities together in a way that suggests that all forms of inequality are alike and 'necessitate similar policies'.⁷⁴ The listing approach in the UK thus has the potential to obscure key differences with respect to inequality on grounds of age, disability, gender reassignment, pregnancy and maternity, race, religion or belief, sex and sexual orientation and indeed the cultural and spatial variation in how these inequalities are experienced.

The second concern identified by Verloo is that single equality instruments perpetuate an individualistic approach that is insufficient to deal effectively with structural disadvantage. To tie this to the discussion about the public sector equality duty, this means that the legal change represented in the Equality Act is not enough: that it must be accompanied by '(...) strategies at the level of structures and institutions'.⁷⁵ In other words, the shift to 'equality' or 'diversity' mainstreaming has to occur in practice at an institutional level as well as in theory. Verloo emphasises that the adoption of 'equality' or 'diversity' cannot simply involve adapting current tools of gender mainstreaming. Rather,⁷⁶

[if] intersectionality is at work in strategies against inequalities, then new and more comprehensive analytical methods are needed and methods of education, training and consultation will have to be rethought.

Although the new EHRC in Britain '(...) requires an approach that acknowledges diversity and intersectionality'⁷⁷ more research is required to determine whether this approach is being applied in practice. However, the fact that current EHRC guidelines for public authorities⁷⁸ do not even make reference to the term 'intersectionality' does little to instil confidence that this is the case. In sum, given the theoretical complexity and contestation surrounding concepts of intersectionality,

⁷² F Mackay & K Bilton, 'Learning from experience: Lessons in mainstreaming equal Opportunities' (University of Edinburgh, Governance of Scotland Forum 2003) 152 <<http://www.scotland.gov.uk/Publications/2003/05/17105/21749>> accessed June 14 2013.

⁷³ M Verloo, 'Multiple Inequalities, Intersectionality and the European Union' (2006) 13 *European Journal of Women's Studies* 211, 214.

⁷⁴ *ibid.*

⁷⁵ *ibid* 215.

⁷⁶ *ibid* 222.

⁷⁷ Bagilhole (n 55) 264.

⁷⁸ Various guidelines are available of the EHRC website: <<http://www.equalityhumanrights.com/advice-and-guidance/public-sector-equality-duty/guidance-on-the-equality-duty/>> accessed May 29 2013.

difference and diversity⁷⁹ it is possible to argue that the attempt to 'blend' equality analysis has happened too quickly, without full appreciation of what these terms mean in theory or practice; the differences between forms of inequality; how they intersect with one another; and what tools are required to ensure that these diverse and intersecting equality concerns are effectively and sophisticatedly integrated into law and policy making. Competition between different forms of inequality is the third potential downfall to a single equality duty that lists multiple protected characteristics. In essence, the concern here is that some equality issues will be paid more attention than others, creating something that resembles a 'hierarchy of inequalities'. This competition, according to Verloo, is '(...) fuelled by the specific nature of current policies' and the political environment at a particular point in time. Gender or disability concerns, for example, could slip down the 'hierarchy' during a moment of heightened political awareness about racial inequality, and thus be less likely to be integrated into law and policy.

5. Conclusion

One of the principal aims of the Equality Act 2010 was to 'harmonise' and 'simplify'⁸⁰ Britain's equality law framework through the creation of a single equality duty covering eight protected characteristics. While the extension of equality protection is welcome and in line with developments elsewhere, this article set out to demonstrate that the background and conceptual underpinning to this fundamental change is far from simple. The new public sector duty not only replaces concerns about inadequate protection with concerns about dilution, vagueness, competition and the adequate recognition of intersectionality, but requires law and policy makers to broaden their equality-sensitive gaze at a time when public resources are incredibly stretched. From the perspective of those concerned with the advancement of gender equality in law and policy making, there is a real risk that gender concerns will be overshadowed or deprioritised and that the ethos of gender mainstreaming will be simultaneously undermined.

To be clear, the movement towards a more 'diversity' or 'equality' focused conception of mainstreaming conception is a step in the right direction. There are few feminists or gender mainstreaming commentators who would deny this, particularly given that 'diversity' or 'equality' mainstreaming is, at least in theory, capable of responding to the needs of *more* women. The point to emphasise, however, is that this new form of mainstreaming will only live up to its conceptual expectations if those responsible for its implementation have a robust understanding of intersectionality, awareness of potential unintended consequences and access to the proper assessment tools. Unfortunately, any hope that this will occur is dampened by the message coming from the top that equality impact assessment is readily dispensable.

⁷⁹ See, e.g., Bacci & Eveline (n 66) 6-8 for an explanation of how the concept of diversity is contested & Verloo (n 73) for an in-depth discussion of intersectionality theory.

⁸⁰ <<http://www.equalityhumanrights.com/legal-and-policy/equality-act/what-is-the-equality-act/>> accessed Aug 5 2013.

Case Comment: *Royal Bank of Scotland v Stuart Hill*

KATHERINE ANDERSON*

1. Introduction and Facts of the Case

*Royal Bank of Scotland v Stuart Hill*¹ (hereinafter 'Hill') is the latest case in which an argument has been made based on the sovereignty and legal position of Orkney and Shetland. Such arguments are not uncommon, either in the courts or in historical record. However, they are based on a misinterpretation of the legal history of the Northern Isles. As such, it is worth considering on what basis the defender made the sovereignty argument and what significance, if any, this has for the position of Orkney and Shetland in Scotland and its legal system. Part of the significance of *Hill* is not the outcome of the case itself but rather the timing of it. Some influential observers, including the islands' MP, have raised the issue of sovereignty for Orkney and Shetland following a vote for Scottish independence.² Previous cases were heard before such a vote was envisaged and could not therefore have considered the legal implications of independence for the Isles.

Hill has, in this and other cases, denied that Shetland is part of the United Kingdom and has thus disputed the jurisdiction of the Scottish courts in the Isles.³ *Hill* stemmed from an attempt to avoid paying a credit card debt by issuing a statutory demand preparatory to a winding-up notice on the Royal Bank of Scotland. Hill demanded payment of £23,029,367.09 from the Bank based on what he regarded as the profits of their fraudulent use of his money. He also created a fee schedule that charged the Bank for letters and phone calls he made to them.⁴ The pursuers disputed that they owed the defender any money, and obtained an interim order preventing the defender from presenting a winding-up petition.⁵ The letter that enclosed the statutory demand was misplaced but RBS did not dispute that it had

* PhD student, School of Law, University of Aberdeen.

¹ *Royal Bank of Scotland v Stuart Hill* 2012 CSOH 110, unreported but available at:

<<http://www.scotcourts.gov.uk/opinions/2012CSOH110.html>> accessed 12 November 2012.

² M Settle, 'Isles May Demand to Remain in UK, Libdem MP warns' *Sunday Herald* (Glasgow, 5 August 2012) <<http://www.heraldscotland.com/politics/referendum-news/isles-may-demand-to-remain-in-uk-libdem-mp-warns.18336259>> accessed 5 November 2012.

³ *Hill* (n 1) [19]; Shetland News 'Hill Locked up as Court Loses Patience' *Shetland News* (Lerwick, 13 July 2012) <<http://www.shetnews.co.uk/news/5105-hill-locked-up-as-court-loses-patience>> accessed 1 November 2012; Shetland News 'Hill Vows to Take Sheriff to Court' *Shetland News* (Lerwick, 19 September 2012) <<http://www.shetnews.co.uk/news/5560-hill-vows-to-take-the-sheriff-to-court>> accessed 1 November 2012. Hill is also chairman of the Shetland and Orkney Udal Law Group, which campaigns for sovereignty of the Isles. See <<http://www.udallaw.com/>> accessed 2 November 12.

⁴ *Hill* (n 1) [8-9].

⁵ *ibid* [10].

been received. The defender therefore argued that the statutory demand had been implicitly accepted. Although the insolvency point was the ostensible purpose of the defender's case, it is probable that he was more motivated by the sovereignty issue that arose from his claim that the court had no jurisdiction over him.⁶

Hill produced a 72-page historical analysis claiming that Shetland is not part of the United Kingdom, an assertion that he based on '(...) the proposition that the Pawning Document granted by King Christian I of Denmark and Norway in 1469 only pledged about 3 or 10 per cent of the land in Shetland.'⁷ This was, according to the defender, a private arrangement relating only to the King's lands in Shetland that left the remainder of the land of the islands in the 'outright ownership of the people of Shetland.'⁸ This view has led the defender to question the jurisdiction of the Scottish legal system in Shetland and the island's position in the UK. Although the defender's arguments specifically related to Shetland, many of the main points in this article are equally pertinent to Orkney, which has a very similar legal history to Shetland. Due to the historical foundation of the defender's claims, this case comment will analyse the history relating to the legal status of the Isles, and demonstrate that the court was right to decide that the defender's position was unfounded.

2. The Judgment

A. The Insolvency Issue

In court, Hill argued that the pursuers accepted the existence of the debt by virtue of the fact that they did not previously object to his claims.⁹ Essentially, the defender attempted to foist an obligation onto the pursuer. Hill argued that there was an implied contract between himself and the bank on the basis that the bank did not object to his terms. Similar arguments have been rejected by the courts in the past. *Carmichael v Black*,¹⁰ for example, considered the right of private companies to clamp cars parked in private car parks following a notice setting out the consequences of improper use, which included fines and clamping. It was concluded that the notice of terms did not conclude any kind of contract as the notice constituted a threat, rather than an offer that could be accepted by choosing to park there.¹¹ It could have been argued that Hill's letter to the bank fitted this example. Another possibility is that the letter had similarities to reliance on unnegotiated terms, such as conditions of carriage. Here, there are usually conditions of carriage that, although not printed on the ticket, are implicitly accepted by choosing to travel. This rule demonstrates that unnegotiated terms can be binding in contract. However, the defender's letter to the

⁶ *ibid* [19].

⁷ *ibid* [19].

⁸ *ibid* [20].

⁹ *ibid* [11].

¹⁰ *Carmichael v Black* 1992 SLT 897.

¹¹ *ibid* [900]. And for discussion of the case see RM White, 'Parking's Fine: The Enforceability of Private Parking Schemes' 2007 JR 1.

bank would not meet the required criteria for such a rule to operate. Lord Pentland thus concluded in *Hill* that '[i]t may (...) be observed that the defender's position is inconsistent with the principle that silence cannot normally be held to import assent'.¹²

An unpaid creditor has a *prima facie* right to wind up a company that has failed to pay its debts.¹³ However, it is required that the party claiming to be a creditor must be owed a valid, legally enforceable, debt.¹⁴ In *Hill*, Lord Pentland interprets this rule to mean that there must be a legitimate basis for the party serving the statutory demand to claim to be a creditor of the company.¹⁵ For reasons discussed above, Lord Pentland concluded that the bank's failure to dispute the debt in response to Hill's letters was insufficient to establish this and, therefore, that there was not '(...) any legitimate basis for [Hill's] claims against the pursuers' in respect of the financial aspect.¹⁶ As a result, Hill could not use section 124(1) of the Insolvency Act 1986 to petition for the winding up of the Royal Bank of Scotland. The defender had attempted to argue that fraud had taken place, however as '[i]t has long been a requirement of our system of pleading that allegations of fraud must be based on clear and specific averments of fact from which an inference of fraud may legitimately be drawn', and Hill had failed to make such averments, his argument was deemed '(...) irrelevant and lacking in specification.'¹⁷

B. Pledging of the Isles to Scotland

The defender's argument that the court had no jurisdiction over him as a resident of Shetland was based on his interpretation of the history of the islands.¹⁸ Given the centrality of this argument to the case, a brief overview of the relevant history is necessary here. In the early 15th century, Orkney and Shetland formed part of the Danish empire which, at this time, included Norway.¹⁹ The Hebrides were part of the Norse empire until 1260 when they were ceded to Scotland in the Treaty of Perth. Under this treaty, annual payments were to be made by Scotland to Norway in return for the islands. By 1468, the Danish King, Christian I, was at war with Sweden. To raise funds, he sought payment by Scotland of arrears owed under the Treaty of Perth. The Scots were unwilling to settle these arrears. However, they sought sovereignty over Orkney which, for all practical purposes, was under the control of a Scot, the Earl of Orkney. A solution was proposed that would marry Christian I of Denmark's daughter, Margaret, to James III of Scotland. Christian agreed to pay a large dowry of 60,000 florins of which he was to pay 10,000 florins immediately and give Orkney in pledge for the remaining 50,000. In 1469, he could only manage 2000

¹² *Hill* (n 1) [16].

¹³ A Keany and P Walton, *Insolvency Law: Corporate and Personal* (2nd ed, Jordan's 2011) [14.8], [14.8.1].

¹⁴ Insolvency Act 1986 section 123 (1); *ibid* [14.8], [14.8.1].

¹⁵ *ibid*; *Hill* (n 1) [15].

¹⁶ *Hill* (n 1) [11].

¹⁷ *ibid* [17].

¹⁸ *ibid* [19].

¹⁹ S Seibert, *Reception and Construction of the Norse Past in Orkney* (Peter Lang 2008) 17; B Crawford, 'The Earldom of Orkney and Lordship of Shetland: a reinterpretation of their pledging to Scotland in 1468-70' (1967-8) 17 *Saga Book of the Viking Society* 156, 161.

florins; Shetland was pledged for the rest. This arrangement benefited Christian significantly: it allowed him to provide his daughter with a dowry at little cost to himself and, in effect, settled a debt which he could not in practice have enforced. The marriage also allowed him access to Scotland's powerful ally, France. In exchange, the Danish King sacrificed Orkney and Shetland, two groups of islands that had little importance to him. This loss only needed to be temporary as the treaty included a right of redemption, although it did not specify a timescale for exercising this right. It was stipulated that if Christian could pay the money he could have the Isles returned to him. Christian commanded his subjects to pay their taxes to the Scottish King until such time as the Isles were redeemed by himself or his successors.²⁰ The defender in *Hill* relied on this possibility of redemption to argue that jurisdiction did not transfer to Scotland and thus that the court did not have jurisdiction over him.

C. Original Intent

The defender appeared to take the view that the agreement between James III and Christian I did not transfer sovereignty.²¹ There has been much debate on this point. Although not expressed in definite terms, the original intent may have been to transfer sovereignty in the Isles to Scotland permanently. It is certainly debatable whether Denmark seriously intended to pay the amounts due. The terms of the exchange seem to have been satisfactory to both parties and it should be noted that, five years after the pledging, Christian raised three times the money for which he had pledged Shetland in order to go to Rome to meet the pope.²² Although it is possible that Christian intended to redeem Shetland, at the very least it can be said that this was not his top priority.²³

D. Extent Pledged

Regardless of what was intended, the defender did not accept that sovereignty was transferred to Scotland, but equally did not seem to believe that it was reserved to Norway and now resides with the people of Shetland.²⁴ Although his documentation for this argument was not reproduced, there seems little to support it either legally or historically. There has been a great deal of debate about the extent and nature of the rights pledged under the treaty. Roberstaad has argued that the pawning comprised the royal sovereign rights, revenue and crown land.²⁵ This seems fairly comprehensive but the documents are unclear on this point. As a result, it has been

²⁰ JV Sigurdsson, 'The Norse Community' in S Imsen (ed), *The Norwegian Domination and the Norse World 1100-1400* (Tapir Academic Press 2010) 35.

²¹ *Hill* (n 1) [19].

²² TK Derry, *A History of Scandinavia* (Allen and Unwin 1979) 79.

²³ BE Crawford, 'The Pledging of the Islands in 1469: The Historical Background' in D Witherington (ed), *Shetland and the Outside World 1469-1969* (OUP 1983) 34.

²⁴ *Hill* (n 1) [20].

²⁵ K Roberstaad 'Udal Law' in D Witherington (ed), *Shetland and the Outside World 1469-1960* (OUP 1983) 56.

argued that it was only the royal estate, and not sovereignty, that was pledged.²⁶ This appeared to be the view taken by the defender. Whereas the 1468 treaty relating to Orkney pledged '(...) all and sundry our lands of the isles of Orkney, with all and sundry rights, services and their just pertinents lawfully belonging to us and our predecessors, the Kings of Norway'²⁷, the Shetland treaty of 1469 built on the Orkney treaty adding '(...) all and sundry our lands of the islands of Schetland, with all and sundry rights and their rightful pertinents whatsoever (...) pertaining or that can pertain in any way to us and our predecessors, kings of Norway.'²⁸ The similarity of, and connection between, these two treaties illustrates that Orkney and Shetland relate to Scotland in the same way. Moreover, the last phrase of the Orkney treaty suggests that there was a transfer of assets or rights which belonged to the King of Norway by right of his office, rather than to him personally. This would seem to amount to sovereignty being passed to Scotland.

Nonetheless, Crawford has argued that it was not *sovereignty* that was pledged, as the parties were concerned not with sovereignty but rather with the right to collect taxes and rents.²⁹ If Crawford's view is correct, the treaty would simply have been about money: James was due money in respect of the dowry which Christian could not pay and he would be able to uplift these payments instead of receiving a lump sum. However, if this had been the case there would likely have been some time-limit governing the period under which the entitlement to uplift rents and taxes would last. Although it would not be unusual to pledge a valuable item such as security for a debt of lesser value, it would not make sense to assign the rents *in lieu* of payment of a lesser amount. A pledge can be redeemed by paying the debt, but the latter would result in a windfall for Scotland. Thus, the way in which the transaction was carried out makes little sense if viewed purely as a cash transaction. The point was for Scotland to gain control of the islands. If the pledging of Shetland had been considered as a cash transaction, a price would have been assigned to the islands that bore some relation to their value. In fact, Shetland was added to the deal at the last minute to make up a shortfall. Orkney was the group in which the Scots were interested - Shetland was closer to Norway and therefore of less interest.

E. Sovereignty

Even if sovereignty was not transferred by the treaty, however, Donaldson has observed that it has long since been transferred by agreement.³⁰ He cites the view of Grant, an international lawyer, who has argued that the islands have been acquired by acquisitive prescription through '(...) peaceful, open, continuous and effective exercise of sovereignty over the territory, coupled with clear and unequivocal

²⁶ Crawford (n 23) 168.

²⁷ J Storer Clouston, *Records of the Earldom of Orkney* (EUP 1914) 56; J Ballantyne and B Smith (eds), *Shetland Documents 1195-1579* (Shetland Islands Council, Shetland Times 1994) document 25.

²⁸ *ibid* (Ballantyne & Smith).

²⁹ Crawford (n 23) 156.

³⁰ G Donaldson, 'Problems of Sovereignty and Law in Orkney and Shetland' in WDH Sellar (ed), *Miscellany Two* (Stair Society 1985) 13, 19.

intention to act as sovereign.’³¹ In the *Hill* case the court accepts this reasoning and concludes on that basis, along with the Crown’s ‘(...) clear and unequivocal intention to act as sovereign’, that jurisdiction has transferred.³²

Contrary to the defender’s argument, it would seem that sovereignty did indeed transfer to Scotland under the treaty. In fact, it is difficult to see how *Hill* could found an idea that sovereignty lies with the people of Shetland. The transfer of the Isles is a transfer of jurisdiction from Norway to Scotland. Such transfers do not always follow the letter of the law and it is often difficult to establish exactly what that law is.³³ The Falkland Islands, another island group with a history of disputed sovereignty, has a more complex history but it is clear that strategic considerations and domestic policy were at least as important as the legal position of the islands and that this is also true with respect to the Scottish Isles.³⁴ However, with respect to the Falklands ‘(...) Britain seems to have consolidated title on the basis of adverse possession and effective occupation’ at the expense of Argentina’s claims.³⁵ This seems to fit with the view held by Grant that regardless of how, when, or even if Scotland acquired sovereignty over Shetland, Scotland *is* now sovereign.

The best-known Northern Isles case is the *St Ninian’s Isle Treasure* case of 1963,³⁶ which considered ownership of treasure found in Shetland. In this case, it was concluded that ‘(...) it is in any event plain that from an early stage the Scottish Parliament assumed the right to legislate for Orkney and Shetland as part of the Kingdom of Scotland.’³⁷ Lord Pentland relies on this case in *Hill* when he states:³⁸

In my opinion, as these statements in the Inner House clearly show, it is now settled that, as a matter of law, Shetland is part of the United Kingdom. The British Crown has the right to exercise sovereignty over the islands. Scots Law applies there and the Scottish courts have territorial jurisdiction there. In the circumstances, I have no difficulty in holding that this court has jurisdiction over the defender in the present action.

Although unequivocal on the sovereignty issue, the *St Ninian’s* case has been criticised by authors, including Smith, for not properly considering the evidence and authorities available on the matter.³⁹ This critique is significant here due to the fact that the court in *Hill* relied on this case. However, it is evident that *St Ninian’s* remains authoritative in spite of the criticism. The Isles fall under Scottish jurisdiction and the court was entitled to judge the insolvency issue over which the *Hill* case was raised.

³¹ JP Grant quoted in G Donaldson, ‘Norse and Scots Law in Shetland’ in *The Shetland Report: A Constitutional Survey* (Nevis Institute 1978) 136, 139.

³² *Hill* (n 1) [23].

³³ W Reisman, ‘The Struggle for the Falklands’ (1983) 93 Yale Law Journal 287, 312.

³⁴ *ibid* [292].

³⁵ *ibid* [296].

³⁶ *Lord Advocate v University of Aberdeen & Budge* 1963 SC 533 (hereinafter ‘*St Ninian’s*’).

³⁷ *ibid* [540].

³⁸ *Hill* (n 1) [28].

³⁹ DL Carey-Miller, ‘St Ninian’s Isle and its Treasure’ in Grant and Sutherland (eds) *Scots Law Tales* (DUP 2010) 113; J Ryder, ‘Udal Law: an Introduction’ 25 (1988) *Northern Studies* 1, 12.

3. Law, Courts and Self-determination

Although there was no provision in the 1469 treaty preserving the Norse laws, from a legal and administrative standpoint the Isles continued much as they always had in the first century after they transferred into Scottish jurisdiction. This is generally the case when territories transfer jurisdiction, as demonstrated by the Scottish position after the wars of independence. The rejection of English authority did not lead to a rejection of its law. In around 1318, when Anglo-Scottish relations were at a low-ebb, the English writer Glanvill was being relied on heavily to create one of the earliest Scottish law texts, *Regiam Majestatem*.⁴⁰ The same phenomenon is seen in Orkney and Shetland and it is suggested here that a transfer of sovereignty will not affect the position of the Scottish legal system in the Isles. The law may never change and, if it does, this will likely be a gradual process as the legal system comes under a different influence. In Scotland, this happened twice: there was a gradual re-orientation towards the Continent after 1314, and then a gradual re-orientation towards England after 1707. Regardless, it has been the accepted approach that indigenous laws remain in place until altered by statute.⁴¹ It is in this context that the Shetland courts were confirming decisions in Bergen as late as the 1530s, although other transactions took place in Edinburgh. This demonstrates that the connections between Norway and Shetland remained strong. It is possible that this confirmation was merely sought as security, as the issue of redeeming the pledge was still live.⁴² It would also appear from the court book of 1602-04 that the local law text, the Lawbook, was still being used in Norway, Iceland and the Faroe Islands.⁴³ This text would be better understood in the Norwegian courts than the Scottish ones.

In 1611, an 'Act Discharging Foreign Laws within Orkney and Shetland' was phrased in a way that suggested that the Isles *should* have been governed under Scots law from the time of the treaty, but that this was not in fact the case.⁴⁴ For example, it states in the Act that it is 'most unlawful' to judge the inhabitants of the Isles by foreign laws and to demand that the 'proper laws of this kingdom' be used.⁴⁵ This was a commonly-held view at this time, and was also espoused by Thomas Craig.⁴⁶ The view seems to be based on the understanding that Scots legal jurisdiction was established in 1469, a fact that would render use of the local laws illegitimate. This Act, however, was not retroactive and did not affect land tenure that was already established. However, it did establish that the legal system was now entirely Scottish.⁴⁷ If it is accepted that sovereignty had transferred to Scotland by this time,

⁴⁰ HL MacQueen, *Common Law and Feudal Society in Medieval Scotland* (EUP 1993) 5; O Robinson, T Fergus and W Gordon, *European Legal History* (3rd ed, Butterworths 2000) [9.6.2].

⁴¹ S Farran, 'Sand, Fish and Sea: A Legal Reflection on Islands – From Orkney to Vanuatu' (2006) 21 *International Journal of Marine and Coastal Law* 389, 393.

⁴² W Thomson, *New History of Orkney* (2nd ed, Kirkwall 2001) 240.

⁴³ B Smith, 'When did Orkney and Shetland become part of Scotland?' (2010) 5 *New Orkney Antiquarian Journal* 52; Ballantyne and Smith (n 27) [40].

⁴⁴ J Ballantyne and B Smith (eds), *Shetland Documents 1580-1611* (Shetland Islands Council Shetland Times 1994) document 528.

⁴⁵ *ibid.*

⁴⁶ T Craig, *Jus Feudale* (Lord Clyde (tr), W Hodge 1934) [1.15.14].

⁴⁷ Ballantyne and Smith (n 44).

143 years after the pledge was made, then the Scottish government was entitled to discharge the indigenous (foreign) laws and confirm the jurisdiction of the Scottish courts. This meant that aspects of the law relating to landholding were still active and considered, under Scots law, as local customary law. Indeed, this is how the law has been treated by the courts.⁴⁸ Although the judgment in *Hill* does not consider the above matters in much depth, it relies on cases such as *St Ninian's*, which did require deeper consideration of these facts.

During the sixteenth century there was a gradual increase in recourse to the Court of Session in disputes over land ownership.⁴⁹ This was presumably perceived as effective given the frequency with which they were used by islanders.⁵⁰ The local tax, *skatt*, was also paid to the Scottish crown, a fact that indicates the islanders' acceptance of Scottish governance but did not establish feudal superiority.⁵¹ However, the people of Shetland appear to have accepted both Scottish law and governance quite readily. The above actions could be characterised as an exercise in self-determination and suggest basic acceptance of the Scottish jurisdiction by islanders at the time. They also support the court's view in *Hill* that '(...) there has been peaceful, open, continuous and effective occupation of Shetland by the British Crown' and that jurisdiction has been established.⁵² This must be the correct view.

4. Conclusion

Jurisdiction over the Northern Isles has passed to Scotland. The defender in *Hill* was held to have failed to establish his argument on this point, in part because even if his assertion was correct, he had evidently accepted that the Scottish legal system had jurisdiction in choosing to use a Scottish legal process.⁵³ However, it might be argued that he did not have any other forum in which to make the argument. In 1968, the Shetland Times asked the Norwegian government, as the original sovereign power in the Isles, for its view on the constitutional position. The response was that '(...) the Orkney and Shetland Islands must be considered properties of the pawnee'.⁵⁴ Norwegian legal opinion was also sought which concluded that '(...) after 500 years, there could never be any question of [Norway] exercising sovereignty rights over the islands'.⁵⁵ This must be accepted as the concluded view of the Norwegian government. However, the islands became part of Denmark with the rest of Norway in 1380 and were not retained by Norway when the union was dissolved. They would therefore remain part of Denmark, as for example, do the Faroes, rather than Norway. There is no reason to suppose that the Danish government pursues a more active role than the Norwegian government. As a result, the defender had no option but to use a Scottish legal process if he intended to make this argument in court.

⁴⁸ Farran (n 41) [389].

⁴⁹ Ballantyne and Smith (n 44) document 250.

⁵⁰ Ballantyne and Smith (n 44) document 533 and 475.

⁵¹ A Menzies, *Conveyancing According to the Law of Scotland* (Bell and Bradfute 1900) 473.

⁵² *Hill* (n 1) [23].

⁵³ *ibid* [20], [21].

⁵⁴ Shetland Times, 22nd March 1968 quoted Smith (n 43) [68].

⁵⁵ *ibid*.

Although he claims that Shetland is sovereign in its own right, the only forum available there is the local sheriff court, run under Scottish jurisdiction.⁵⁶

Scottish courts have accepted jurisdiction in the Isles for centuries and this has been examined in a number of cases.⁵⁷ Orkney and Shetland are part of the Sheriffdom of Grampian, Highlands and Islands. Appeals are made in the usual manner.⁵⁸ The historical evidence tends to support Lord Pentland's decision in this case. The political position is that neither Norway nor Denmark seeks jurisdiction over the islands, where the UK and Scots law regularly exercise such jurisdiction. We must conclude that the defender's argument regarding jurisdiction had no likelihood of success. However, the argument has been made before and, given the arguments made with regard to the status of the islands if Scotland became independent, it is probable that it will be raised again. Hill may appeal this case, and a similar case in which he has made the sovereignty argument in the Sheriff court is being appealed to the High Court.⁵⁹

Hill suggests the direction that would be followed should any legal action be raised regarding Shetland's status on independence. The case supports the view that the Isles are within the territorial jurisdiction of the Scottish courts and part of the UK.⁶⁰ As such they remain part of Scotland and would automatically leave the United Kingdom with the rest of Scotland unless a separate arrangement was made. Shetland's MSP, Tavish Scott, has recently suggested that Shetland would benefit from further powers being devolved.⁶¹ Although he does not question the Northern Isles' legal status, he is making an argument as to how he believes the Northern Isles could turn the current Scottish political situation to their advantage. It is because of issues like this being raised that a clear and unambiguous statement of the constitutional position of the Isles and their status within the Scottish judicial system is useful. *Hill* provides that statement.

⁵⁶ *Hill* (n 1) [19].

⁵⁷ See for example *Lord Advocate v University of Aberdeen & Budge* 1963 SC 533 or *Sinclair v Hawick* (1624) Mor 16393).

⁵⁸ Scottish Courts Website <<http://www.scotcourts.gov.uk/sheriff/gazetteer.asp>> accessed 2 November 2012.

⁵⁹ The Sovereign State of Forvik <<http://www.forvik.com/news>> accessed 15 February 2013.

⁶⁰ *Hill* (n 1) [22] - [23].

⁶¹ BBC News 'Lib Dem Tavish Scott Calls for Home Rule in the Northern Isles' *BBC News* (London, 16 March 2013) <<http://www.bbc.co.uk/news/uk-scotland-scotland-politics-21815296>> accessed 18 March 2013.

Book Review

Child Abduction within the European Union
Katarina Trimmings
(Hart Publishing, Oxford 2013, ISBN: 9781849463973)

Cross-border parental child abduction is a global problem that affects thousands of children every year. The key instrument designed to protect these children from the harm caused to them by their abduction is the Hague Convention on the Civil Aspects of International Child Abduction (1980) ('Hague Abduction Convention'). The objective of this Convention is not to adjudicate on custody issues but rather '(...) to secure the prompt return of children wrongfully removed or retained in any Contracting State.' The rationale underlying the Convention's approach is two-fold. In addition to acting as a deterrent to abduction, it is thought to be better for the child to be returned to their normal routine as quickly as possible and to be reunited with the left-behind parent. The more quickly that the child can be returned to their state of origin, the less likely it is that the child will become settled in the state of refuge. However, if the child does become settled due, for example, to delays in the judicial process, the child's 'new' habitual residence could be used as an argument by the abducting parent for the non-return of the child. This latter point caused a deep division between Member States during the negotiations for the Brussels *Ibis* Regulation; between those who perceived a regional EU instrument dealing with child abduction as necessary in order to prevent this perceived misuse of the Hague Abduction Convention and those who felt that the Hague Abduction Convention worked well as it stood.

As the author of this book explains, pro-reformists within the EU argued that claims by the abducting parent that the child had settled in the state of refuge and therefore could not be returned, was a misuse of Article 13(1)(b) of the Hague Abduction Convention. Article 13(1)(b) states that the child should not be returned to the state of origin if '(...) there is a grave risk that his or her return would expose the child to physical or psychological harm or otherwise place the child in an intolerable situation.' Despite counter arguments that additional legislation would lead to '(...) confusion for parties, practitioners and judges', the above-noted concern led to the creation of a separate intra-EU child abduction regime in the form of the Brussels *Ibis* Regulation. It is the evolution of this separate intra-EU child abduction regime that this book seeks to critically evaluate. The author questions whether the Brussels *Ibis* Regulation has added anything of value to the law governing child abduction, arguing convincingly that there was no real legal need for either the involvement of the EU in the area of child abduction, or for the tightening of the 1980 Hague Child Abduction Convention return mechanism, under which the Regulation is now able to 'trump' the Convention in certain circumstances.

Chapters two and three therefore provide a chronological account of the development of the Brussels *Ibis* Regulation. The author is highly critical of the way in which negotiations were conducted during the revision of the Regulation. Of particular concern, which should sound alarm bells with anyone interested in EU law-making, was the apparent lack of any empirical research by the European Commission to justify the need for a European Union solution to child abduction, the lack of expert involvement and the Commission's reluctance to take objections raised by Member States on board. She notes that provisions that appeared during the negotiations, that could have been of actual benefit, (for example the provision of guidelines for special cases, such as those involving spousal violence) were inexplicably omitted from the final Regulation. An interesting comparison is made between the many EU principles of good legislative drafting and the process that was followed during the drafting of the child abduction provisions in Brussels *Ibis* Regulation. The suggestion is that there is a gap between the high standards that the EU purports to uphold with respect to the legislative drafting, and the practical reality.

In chapter 4, the author comprehensively challenges the argument made by the pro-reformists that there was a misuse of Article 13(1)(b) of the Hague Abduction Convention, and considers the textual evidence that would have been available to the Commission prior to the adoption of the Brussels *Ibis* Regulation. After analysing a wide range of material, including Hague Conference Special Commission documents, conclusions of conferences and academic literature, she draws the stark conclusion that there was no evidence to justify the need for the separate intra-EU child abduction regime and, further, that the evidence actually points to an efficient and functioning Hague Abduction Convention.

Chapter 5 contains the results of the author's own comprehensive empirical research on the practical operation of the Brussels *Ibis* Regulation. Her detailed statistical analysis of the return applications confirms widely held assumptions, including that the majority of the abducting parents are mothers and that more than half of the abductors are 'returning home' to the Member State of their nationality. Of particular interest is that the Regulation does appear to have been successful in reducing the use of the 'grave risk of harm', Article 13(1)(b) defence, which as previously mentioned was one of the issues that the Regulation sought to address. However the author is keen to point out that there also appears to be evidence of avoidance with respect to the use of Article 13, combined with a concurrent increase in the use of non-Article 13 defences, that would go some way to explaining this reduction.

Important questions are posed in chapter 6 regarding the judicial and administrative qualities in 'new' EU Member States in Eastern Europe, in particular the question of why such states make far fewer return orders than the EU average. The author's explanation for this is illuminating. Basing her conclusions on evidence from her interviews with representatives from the Central Authorities of Slovakia and the Czech Republic, she cites fundamental historical differences between Central and Eastern European countries as the underlying cause, arguing that the transformation in Eastern European countries, from communist to democratic systems of government, has been a slow process. Moreover, the author suggests that

members of the judiciary in these countries still lack experience in applying international law and are struggling to '(...) overcome methodological problems of post-communist legal reasoning such as extreme formalism and textual positivism.' The author notes that there is evidence to support the claim that the judiciary in Slovakia and the Czech Republic remain wary of returning an abducted child to a 'capitalist' state where there are doubts over whether the best interests of the child, in their view, can be protected. This failure to apply the rule of law is obviously disturbing when one considers that a fundamental principle of the European Union is mutual trust between Member States.

Although this book is highly critical of the Brussels IIbis Regulation the author does not shy away from praise where she considers it due. In chapter 7, she identifies an increasingly child-centred approach within Europe, (England and Wales and Ireland in particular) that has clearly developed as a result of the Regulation's obligation to hear and pay heed to the views of the child. Through her excellent comparative analysis of the current approaches towards the defence of the child's objections found within case-law from both EU jurisdictions and non-EU jurisdictions, the author is able to show that this child-centred approach is not evident outside of the EU and is to be attributed to the Regulation's principle of automatic inquiry into the views of the child. Interestingly, the author draws attention to the different approaches taken by EU Member States to the issue surrounding the age at which the child must be given the opportunity to be heard, highlighting the disparity that still exists in this area. While the views of children as young as six are being taken into consideration by some courts, others have determined that this is too young.

The author's critical assessment of the development of the recent intra-EU child abduction regime therefore raises serious questions about law-making within the EU that should not be ignored. Her recommendations for amending the Regulation, which include the need to secure the safety of not only the child but also of the mother in cases of spousal violence are valid, but, as she argues, can only be successful if the Central Authority of the Member State is able to guarantee their enforcement. These are recommendations that the European Commission would do well to listen to.

All in all, this book is written in a clear, logical and succinct style, offering essential reading for academic study of The Hague Abduction Convention or Brussels IIbis Regulation, as well as for any reader who wishes to gain an insight into EU law-making. In a wider socio-political context, the fact that parental child abduction affects so many children around the world means that this book may attract readers from many different fields with an interest in how society deals with this extremely emotive issue.

Jayne Holliday

May 2013

The Aberdeen Law Project AGM 2013

MALCOLM M. COMBE*

This text is an amended and expanded reflection on a talk delivered at the Aberdeen Law Project AGM on Monday 11 March 2013.

Lectures often make use of pictorial aids, which can make subsequent textual reproduction of those lectures a little tricky. In my address to the Aberdeen Law Project AGM, five pictures were referred to. The imagination of the reader (possibly with the aid of an internet search-engine or a good, old-fashioned encyclopaedia) will be called upon here to visualise three of the five pictures, namely: a horse and cart; the Forth Bridge; and a record (as in a vinyl LP).

For ease of reference, two images can be safely reproduced here.



So, how do you knit together a lecture on *pro bono publico* student legal advice from those five pictures?

The Aberdeen Law Project logo can be explained thus. It has something of a dual identity, hence the Latin and English names. Clearly, the English name is not a direct translation. So why *Casus Omissus*? As explained by Ryan Whelan in an article in the first Aberdeen Student Law Review,¹ the Project '(...) is concerned with the gaps within our legal system.' Unmet legal need. Access to justice.² You know the buzz phraseology, but the Project tries to make a difference.

* Lecturer, School of Law, University of Aberdeen.

¹ RT Whelan, 'Casus Omissus – The Aberdeen Law Project' (2010) 1 Aberdeen Student Law Review 137, 137.

² See generally H Genn and A Paterson, *Paths to Justice: Scotland* (Hart Publishing 2001) and more recently A Paterson, *Lawyers and the Public Good: Democracy in Action? The Hamlyn Lectures, 62nd Series* (Cambridge University Press 2011).

The University of Aberdeen logo can be explained in heraldic terms (and is so explained in an excellent Gaudie article,³ the Gaudie being another fine student institution at Aberdeen). As interesting as that is, heraldry is not the subject of this discussion. Rather, the logo can be taken to symbolise education, i.e. the University of Aberdeen is an educational institution. It has been since 1495. Sure, other things happen at the university, but education comes first, does it not? The University of Aberdeen's registration with Scottish charity regulator OSCR might imply so,⁴ where it notes the following purposes: 'The advancement of education, The advancement of health, The advancement of citizenship or community development, The advancement of the arts, heritage, culture or science.'

The horse and cart is a rather dated form of transport. It lives on in the English phrase 'putting the cart before the horse', aka running before you can walk.

The Forth Bridge is a wonder of cantilever engineering. Its intricate design makes (or used to make) painting and maintenance a never-ending task. That task lives on in the phrase 'like painting the Forth Bridge', used for an endless, repetitious task. Sisyphus, re-start your uphill struggle.

The vinyl record may not mean very much to Generation Y, but it used to be the prevalent medium for music. When an LP gets stuck, repetition happens. When an LP gets stuck, repetition happens. (See what I did there?)

Putting the cart before the horse. Painting the Forth Bridge. Sounding like a stuck record. Are there up-to-date equivalents for these three phrases? Maybe, but not many were proffered when I asked those assembled at the AGM. All three of those phrases could do with a makeover to bring them up to the present day. Their time has passed.

Pro bono publico legal activity has a long pedigree, with Scotland being able to look back to the 1424 Poor's Roll. The late Lord Bingham described that as '(...) the world's first statutory authority on legal aid for the poor',⁵ so the long *pro bono publico* pedigree is clear, but its time has certainly not passed. In any legal system where legal advice is not free to all at point of need, unmet legal need can be a problem.

Scotland was not an exception in the fifteenth century and it is not an exception now. With increasing strain on state-funded legal aid, many different resources are being called upon to take up the slack. Broad potential solutions include new money,⁶ new

³ S Raybaud, 'The Aberdeen University shield' (27 November 2012)

<<http://www.thegaudie.co.uk/features/the-angus-diaries/2012/11/the-aberdeen-university-shield/>> accessed July 1 2013. The Gaudie is the University of Aberdeen's student newspaper.

⁴ <<http://www.oscr.org.uk/>> accessed 1 July 2013.

⁵ T Bingham, *The Rule of Law* (Penguin Books 2011) 87. The history was expertly explained in a previous address to the Aberdeen Law Project AGM by Gary Allan QC, as digested in this journal: G Allan, 'Casus Omissus: The Aberdeen Law Project Inaugural Lecture, 07 March 2011' (2011) 2 Aberdeen Student Law Review 12.

⁶ This is particularly topical in relation to criminal legal aid in Scotland, following the Scottish Civil Justice Council and Criminal Legal Assistance Act 2013 that paves the way for a compulsory

forms of delivery⁷ or new providers. When considering all these bridging resources and their respective sub-categories, Paterson noted that student law clinics '(...) have a real part to play' as a provider in our emerging legal order.⁸ So what are these student law clinics? As many familiar with the Aberdeen Law Project will know, they are free advice centres staffed by (as yet) unqualified solicitors offering support to those who need a lawyer but cannot obtain legal advice elsewhere. Normally, this is because such clients (if that is the correct term) fall in the gap between those ineligible for legal aid and those able to afford a lawyer, hence the need for students to step forward and offer their services *pro bono publico* (normally shortened to *pro bono*) – for the good of the public.

Students can have a role to play, as I have argued before in a blog post⁹ and as others have argued in more learned terms.¹⁰ It is important that any initiative they have is fostered and an environment is provided for that. That is where I come to sound like a stuck record. In my brief address to the 2012 AGM I mentioned how, despite occasional harrumphs, I was (generally) delighted to help any students knocking on my door for a bit of assistance with a case they are advising on. I said the same again in 2013. I may say the same in 2014. When an LP gets stuck, repetition happens.

Why do students have a role to play? Unmet legal need, the *casus omissus*, is not going away. In fact, it is a bit like painting the Forth Bridge. You think you are about 3/4 done, and it is time to start again. Again. That is why student law clinics and those who assist student volunteers have a role: today, tomorrow and the day after tomorrow. You get the drift.

Should these students be helping? After all, or so the argument goes, they are not ready, or at least not rubber-stamped. Some are not even half-way through an undergraduate degree. Is it all a bit like putting the cart before the horse? Maybe, but so be it.¹¹ Even with such reservations, with suitable peer-observation and supervision a law student can make a real difference to someone with an access to

contribution to legal costs when earnings are above a certain threshold. For a civil law perspective (from England & Wales), see e.g. Lord Jackson, *Review of Civil Litigation Costs: Final Report* (2010) <<http://www.judiciary.gov.uk/NR/rdonlyres/8EB9F3F3-9C4A-4139-8A93-56F09672EB6A/0/jacksonfinalreport140110.pdf>> accessed 1 July 2013.

⁷ Perhaps through increased usage of technology to mass-produce legal services (see especially R Susskind, *The End of Lawyers?: Rethinking the nature of legal services* (OUP 2010)) or through innovations like telephone advice lines or websites (such as <<http://www.adviceguide.org.uk/scotland.htm>> accessed 1 July 2013).

⁸ Paterson, *Lawyers and the Public Good* (n 2) 116. See also the discussion in MM Combe, 'Review of Alan Paterson, "Lawyers and the Public Good: Democracy in Action?" (2012) 16 Edinburgh Law Review 466.

⁹ MM Combe, 'Is there a lawyer in the house? Or a law student?' (7 August 2012) <<http://basedrones.wordpress.com/2012/08/07/is-there-a-lawyer-in-the-house-or-a-law-student/>> accessed 1 July 2013.

¹⁰ S Wizner, 'The Law School Clinic: Legal Education in the Interests of Justice' (2002) 70 Fordham Law Review 1929, also at Faculty Scholarship Series, Paper 1843 <http://digitalcommons.law.yale.edu/fss_papers/1843> accessed 1 July 2013.

¹¹ For discussion, see S Zeidman, 'Sacrificial Lambs or the Chosen Few?: The Impact of Student Defenders on the Rights of the Accused' (1996) 62 Brooklyn Law Review 853.

justice issue.¹² Did I mention access to justice issues are not magically disappearing in 2013?

In the process of making a difference, students will learn so many valuable insights about how the law works, clients and perhaps even about themselves. Thus we return to the one picture remaining. Education. Whether as part of a credit-bearing course or as part of a co-curricular activity, *pro bono publico* student legal activity can make a difference to students.¹³

And to the community.

And even to the academic staff at universities.

Why staff? Well, I recently circulated a survey to students involved with the Aberdeen Law Project. My main concern was to gauge appetite for a credit-bearing course at the educational institution at which the students and I are based, but I took the opportunity to ask (almost in passing) why students got involved with the Aberdeen Law Project. Some truly inspirational replies have come back. I am absolutely delighted to be involved with this particular student activity. The students make a difference and are an inspiration to me and hopefully to anyone who attended the AGM or to those reading this note.

Finally, you now know how to get from a horse and cart, to the Forth Bridge, to a record, with a spattering of education and student voluntary work. Perhaps that too will be inspirational.

¹² G Critchlow, 'Professional Responsibility, Student Practice, and the Clinical Teacher's Duty to Intervene' (1991) 26 *Gonzaga Law Review* 415.

¹³ See generally K Kerrigan and V Murray, *A student guide to clinical legal education and pro bono* (Palgrave Macmillan 2011); C Sylvester, 'Bridging the Gap? The Effect of Pro Bond Initiatives on Clinical Legal Education in the UK' (2003) 3 *International Journal of Clinical Legal Education* 29 and, on the layered benefits of the clinical experience, D Nicolson, 'Education, education, education': Legal, moral and clinical' (2008) 42:2 *The Law Teacher* 145. See also E Chemerinsky, 'Why Not Clinical Education', where he notes: 'There is no better way to prepare students to be lawyers than for them to participate in clinical education' (2009) 16 *Clinical Law Review* 35, 35.

Guidelines for Contributors

Submissions are welcomed from students and graduates of the University of Aberdeen.

Several types of articles will be considered:

- Case Notes (500 to 1,500 words on a recent judicial development)
- Book Reviews (500 to 1,500 words on a recent publication)
- Essays and Short Articles (1,500 to 4,000 words)
- Long Articles (4,000 to 10,000 words)

The limits specified are for general guidance only and will not be strictly upheld.

Every piece of work is reviewed anonymously by the student Editorial Board, before being sent for Peer Review. Submissions may be accepted outright, accepted subject to modification, or rejected. Constructive comments will be sent to the author.

Published articles will be compiled in an annual issue which will be available online and in law libraries throughout the country.

All submissions must conform to House Style and to the Oxford Standard for Citation of Legal Authorities (OSCOLA).

Articles must provide a critical analysis of a particular area of law, publication or judicial decision. Comment should be original, relevant and aim to make an interesting contribution to the academic debate.

For further information on submitting an article, subscribing to the journal or becoming involved in a supportive capacity, please contact us via the details on our website: www.abdn.ac.uk/law/aslr.

We would be delighted to hear from you.

This publication has been generously sponsored by Stronachs LLP.

All those involved with the Aberdeen Student Law Review would like to thank Stronachs for their support.

The logo for Stronachs, featuring the word "Stronachs" in a white, serif font, centered within a solid red rectangular background.

Stronachs

For further information please contact

Stronachs LLP
34 Albyn Place
Aberdeen
AB10 1FW

info@stronachs.com
01224 845845

www.stronachs.com