Northern Ireland Legal Quarterly

Volume 71 Number 3

Autumn 2020

CHIEF EDITOR

DR MARK FLEAR

NILQ

Chief Editor

Mark Flear

Co-Editor

Heather Conway

Editorial Board

Yassin Brunger

David Capper

Clayton Ó Néill

Paulina Wilson

Production Editor

Marie Selwood

Northern Ireland Legal Quarterly Volume 71 Number 3

Contents

ARTICLES

Patents, access to medicines and COVID-19: the role of compulsory and government-use licensing in Ireland
Aisling McMahon 331

Abortion activism, legal change, and taking feminist law work seriously
Máiréad Enright, Kathryn McNeilly and Fiona de Londras 359

The 'one man company' after *Patel v Mirza*: attribution and illegality in Singularis Holdings v Daiwa Capital Markets
James C Fisher 387

Duress and loss of control: fear and anger in excusatory defences
Glenys Williams 401

Recalibrating the governance of remedial clauses in contract law
Moshood Abdussalam 429

Mediation in French administrative courts: what lessons for administrative justice?
Sophie Boyron 457

The Singapore Convention: a solution in search of a problem?
Brian Clark and Tania Sourdin 481

NOTES AND COMMENTARIES

The Supreme Court's judgment in *Adams* and the missing step of statutory construction
David Horvath-Franco 503

Procurement and commissioning during COVID-19: reflections and (early) lessons
Albert Sanchez-Graells 523

Pre-charge identification of a minor and Article 14 of the ECHR: Judgment *In the Matter of an Application by JKL (A Minor)*
Faith Gordon 531

ISSN 2514-4936 (online) 0029-3105 (print)

NILQ 71(3): 331–358

Patents, access to health and COVID-19 – The role of compulsory and government-use licensing in Ireland

AISLING MCMAHON*

Maynooth University

Abstract

As the race for effective vaccines and treatments for COVID-19 continues, attention must turn to how such health-technologies will be accessed globally once developed. Patents play a significant role in this context because they give the patent-holder the right to stop others using patented inventions. Patents are available on diagnostics, medicines and vaccines and could form significant access obstacles for COVID-19. Moreover, whilst many patent-holders may be willing to license health-technologies reasonably, others may not. Therefore, it is imperative that national governments ensure effective avenues exist to intervene with patent-holder discretion via compulsory licensing. This article focuses on the legal framework applicable in Ireland for such compulsory licensing interventions, interrogating the effectiveness of the current framework in alleviating access issues posed by patents for COVID-19. It demonstrates how the current framework could be reformed to make it more effective in tempering patent-holder control, where needed, whilst remaining in compliance with Ireland's international obligations.

Keywords: patents; COVID-19; compulsory licensing; government-use licence; service of the state; access to medicines.

Introduction

The race to secure effective vaccines and treatments for COVID-19 continues at pace.[1] However, as we get closer to finding effective vaccines and treatments for COVID-19, attention has turned to how they will be accessed once developed, by whom first and on what terms.[2]

* Assistant Professor in Law, Maynooth University. Email: aisling.mcmahon@mu.ie. The author would like to thank the anonymous peer reviewers for their very helpful comments. This article discusses the legal position and developments as applicable at the time of writing in September 2020.

1 Ewen Callaway, 'The race for coronavirus vaccines' (2020) 580 Nature 576; Emiliano Rodríguez Mega, 'Latin American scientists join the coronavirus vaccine race: "No one's coming to rescue us"' (2020) 582 Nature 470; Ewan Callaway, 'The unequal scramble for coronavirus vaccines – by the numbers' (2020) 584 Nature 506.

2 These include discussions in the vaccine context around COVAX a global initiative seeking to pool vaccine procurement for COVID-19 that is co-led by the Coalition for Epidemic Preparedness Innovations, Gavi, the Vaccine Alliance, and the WHO. See discussions: 'COVAX initiative: the COVID-19 vaccine Global Access Facility' (25 August 2020) <www.healtheuropa.eu/covax-initiative-the-covid-19-vaccine-global-access-facility/102337>; WHO, '172 countries and multiple candidate vaccines engaged in COVID-19 vaccine Global Access Facility' (News Release, 24 August 2020) <www.who.int/news/item/24-08-2020-172-countries-and-multiple-candidate-vaccines-engaged-in-covid-19-vaccine-global-access-facility>. See also discussions of whether there will be enough supplies of the vaccine (once developed): Roxanne Khamsi, 'If a coronavirus vaccine arrives, can the world make enough?' (2020) 580 Nature 578.

In this context, intellectual property rights and particularly patents play a significant but sometimes overlooked role.[3] A patent is an intellectual property right which allows the patent-holder to exclude others from using the invention under patent without the patent-holder's permission (licence). This in turn means that the patent-holder can dictate many aspects of how a patented invention is provided, including, at what price and to whom.[4] Where the patented invention is a health-related technology such as a vaccine, medicine or element(s) of a diagnostic, how patents are used has significant implications for healthcare because access to such technologies is often dependent on how patent-holders choose to license them. The stakes are heightened in the global pandemic context, where access to effective vaccines, medicines and diagnostics is key to saving lives and to controlling and limiting the spread of the virus. Moreover, alongside such implications for health, access to such Covid-19 health-technologies is also vital to alleviate the attendant devastation the pandemic continues to bring in terms of its effect on society and the economy more generally.[5]

Whilst many patent-holders have shown willingness to offer favourable licensing terms for COVID-19 health-related technologies,[6] there is no legal requirement for rights-holders to do so, and others in future may not. This in turn could limit access to and supply of patented health-technologies including vaccines, medicines and diagnostics for COVID-19.

Against this backdrop it is important to consider what avenues are available to national governments to intervene with how patent-holders license patented health-technologies for COVID-19 in cases where patent-holders refuse to license such technologies, or refuse to license them on reasonable terms. This article examines this issue, focusing specifically on the Irish jurisdiction as a case study and on the legal avenues for compulsory licensing and mechanisms for licensing for service of the state applicable in Ireland.[7]

Compulsory licensing allows states to authorise a third-party/government to use a patented technology without the patent-holder's consent. Whilst provisions for licensing for the service of the state (so-called 'government-use provisions'), where applicable within national laws, allow the government to license a patented technology (without the patent holder's consent) for use for service of that state. Both mechanisms, where authorised, facilitate a third party/the state using a patented invention without patent-

3 Other forms of intellectual property rights may also be relevant to medical technologies, discussed in part one below. In the vaccine context, the vaccine production process may often be protected by trade secret protection, and access to such trade secret information and the know-how of vaccine production can prove additional obstacles for a generic company to recreate a vaccine. See Sarah Eve Crager, 'Improving global access to new vaccines: intellectual property, technology transfer, and regulatory pathways' (2014) 104(11) American Journal of Public Health e85–e91.

4 Aisling McMahon, 'Biotechnology, health and patents as private governance tools: the good, the bad and the potential for ugly?' (2020) 18(3) Intellectual Property Quarterly 161–179.

5 In Ireland, the GDP is expected to decrease by approximately 13% in 2020 due to economic disruptions caused by the COVID-19 crisis. See Kelly C De Bruin, Eoin Monaghan, Aykut Mert Yakut, 'The environmental and economic impacts of the COVID-19 crisis on the Irish economy: an application of the I3E model' (ESRI Research Series 106, July 2020) <https://doi.org/10.26504/rs106>; the UN predicts that 40–60 million people globally will be pushed into extreme poverty due to economic effects of COVID-19 crisis: UN Development *Programme, Brief 2: Putting the UN Framework for Socio-Economic Response to Covid-19 into Action: Insights* (June 2020) <www.undp.org/content/undp/en/home/coronavirus/socio-economic-impact-of-covid-19.html>.

6 See discussion in Aisling McMahon, 'Global equitable access to vaccines, medicines and diagnostics for Covid-19: The role of patents as private governance' (2020) Journal of Medical Ethics (forthcoming).

7 The article uses the term Ireland to describe laws applicable within the Republic of Ireland.

holder consent where needed and could be used where public health requires it, for example, where necessary within the COVID-19 context.

Accordingly, the article argues that it is incumbent upon the Irish government, and other national governments, to re-evaluate the current operation of such licensing mechanisms to ensure these mechanisms are as effective as possible for use within the COVID-19 context should they be required. Indeed, other jurisdictions including Germany,[8] France[9] and Canada[10] have already taken steps to reform or tailor existing compulsory licensing frameworks to ensure effective avenues exists to temper patent-holder control where needed for COVID-19.[11] Moreover, a broader shift within the discourse is evident around how compulsory licensing provisions are being discussed in the COVID-19 context: as where previously such provisions were viewed as exceptional measures sitting at the margins of patent law discourse, COVID-19 has brought such provisions under the spotlight as viable and necessary avenues for states to use to ameliorate access issues posed by patents on COVID-19 health-technologies. Yet, there is currently a dearth of literature on the operation of such provisions generally under Irish law[12] and, particularly, a notable gap on work assessing how such provisions might apply within the broader health context in Ireland, including for COVID-19.

This article fills this gap providing an overview of how these provisions operate in Ireland and offering the first comprehensive analysis of how such Irish provisions would likely apply if needed for COVID-19 including the shortcomings evident. In doing so, the article also puts forward novel reform proposals in relation to the Irish framework to make such licensing mechanisms more effective in alleviating potential access issues posed by patents on health-technologies for COVID-19 and within the health context more generally. Importantly, the article argues that current provisions within Irish law could in theory be interpreted as allowing the grant of compulsory licences or government-use licences in the COVID-19 context in *some circumstances*, but that shortcomings remain within the current framework. The amendments suggested to the domestic legal framework are aimed at increasing the effectiveness of the system, both in:

---

8 See discussion in Jennifer Enmon and Grant Shoebridge,'COVID-19 – patent rights in the time of a pandemic' (*Lexology*, 27 August 2020) <www.lexology.com/library/detail.aspx?g=a992bce6-c3ab-41f4-9e30-1fa46861d6f1>.

9 Ibid.

10 Ed Silverman, 'A Canadian bill would make it easier to issue compulsory licenses for Covid-19 products' (*Stat News*, 25 March 2020) <www.statnews.com/pharmalot/2020/03/25/canada-compulsory-license-coronavirus-covid19>.

11 See general discussion of such moves in Adam Houldsworth, 'The key covid-19 compulsory licensing developments so far' (IAM, 7 April 2020) <www.iam-media.com/coronavirus/the-key-covid-19-compulsory-licensing-developments-so-far>.

12 For a discussion of such mechanisms generally, see Robert Clark, Shane Smyth and Niamh Hall, *Intellectual Property Law in Ireland* (4th edn, Roundhall 2016) chapter 8, 'Voluntary and compulsory licenses'; European Patent Academy, 'Compulsory licensing in Europe: a country-by-country overview' (European Patent Office 2018), 61–63. For a brief discussion of the potential application of compulsory licenses for COVID-19 in Ireland, see Sophie Delaney, 'Compulsory licences on the horizon for drugs and equipment?' (*William Fry*, 27 March 2020) <www.williamfry.com/newsandinsights/news-article/2020/03/27/compulsory-licences-on-the-horizon-for-drugs-and-equipment>; Donal M Kelly, 'COVID-19: the impact on IP law and practice in Ireland' (*IP Stars*, 5 May 2020) <www.ipstars.com/NewsAndAnalysis/COVID-19-The-impact-on-IP-law-and-practice-in-Ireland/Index/5589>; Samantha Silver and Lindsay MacLean, 'COVID-19: vaccine development and compulsory licensing' (*Kennedys Law*, 14 May 2020) <www.kennedyslaw.com/thought-leadership/article/covid-19-vaccine-development-and-compulsory-licensing>; Aisling McMahon, 'How patents will affect pandemic vaccines and treatments' (*RTE Brainstorm*, 3 July 2020) <https://www.rte.ie/brainstorm/2020/0702/1150969-patents-coronavirus-vaccines-medicine-ireland>.

1) expanding/clarifying the scope for the use of such provisions for COVID-19; and 2) explicitly acknowledging within the legislation that such provisions can be used within the public health context thereby encouraging the use of such provisions where needed in Ireland.

Notably, alongside national laws, Ireland has obligations under EU law and international laws applicable in the compulsory licensing context. This article provides an overview of how these differing levels of obligations apply to Ireland's domestic framework, and the extent of reforms possible in this context, if Ireland is to remain compliant with such international legal obligations.[13]

Moreover, although the arguments focus on the Republic of Ireland, such arguments have broader significance given that the obligations imposed by the Agreement on Trade-Related Aspects of Intellectual Property Rights 1995 (TRIPS Agreement) apply in all World Trade Organization (WTO) contracting states including EU states and the EU in its own right, given that the EU is a signatory to the TRIPS Agreement.[14] The UK is also a signatory of the TRIPS Agreement and has participated in other WTO instruments, including the Amendment to the TRIPS Agreement in 2005,[15] by virtue of the EU's signing such agreements which bound all EU states (including the UK) at the time of signature. The UK has signalled its intent to remain party to such arrangements/agreements post-Brexit.[16] Accordingly, the WTO legal framework discussed in this article, which defines the parameters within which Ireland can reform its laws and remain compliant with its international law obligations, is applicable in all other WTO states including all EU states and the UK. Thus, this discussion is of broader relevance to any WTO state contemplating reform of its compulsory licensing laws in a manner which remains in compliance with WTO obligations. The obstacles posed by EU law to compulsory licensing use are also of relevance to all current EU states.

Furthermore, the legislative framework applicable in Ireland for compulsory licensing is broadly similar to the framework applicable within the UK under its Patents Act 1977, as amended, as many of the provisions applicable under Irish law were originally drawn from UK law.[17] Moreover, both jurisdictions have licensing for government/crown-use

---

13 Failing to comply with such international obligations under the Agreement on Trade Related Aspects of Intellectual Property Law (TRIPS Agreement) could render Ireland liable to the WTO dispute settlement mechanism and the possibility of trade sanctions for non-compliance. See general discussion on TRIPS Compliance in Edward Lee, 'Measuring TRIPS compliance and defiance: the WTO Compliance Scorecard' 18 (2011) Journal of Intellectual Property Law 401. Moreover, as the EU has signed the TRIPS Agreement, the Agreement has legal effects within the EU legal order, see Justine Pila and Paul Torremans, *European Intellectual Property Law* (Oxford University Press 2016) 35 and 61.

14 The EU signed the TRIPS Agreement on behalf of EU states as it was deemed to be within the competence of the EU: see Council Decision (of 22 December 1994) concerning the conclusion on behalf of the European Community, as regards matters within its competence, of the agreements reached in the Uruguay Round multilateral negotiations (1986–1994); OJ L 336 of 23/12/1994, page 1.

15 WT/L/641, Amendment of the TRIPS Agreement (8 December 2005).

16 The UK has indicated that during the Brexit transition period it will continue to be treated as a member state of the EU for the purpose of such instruments, and after this transition period the UK has confirmed its continued acceptance of such agreements/arrangements, see 'The United Kingdom's withdrawal from the European Union – Communication from the UK' WT/GC/206 <https://docs.wto.org/dol2fe/Pages/SS/directdoc.aspx?filename=q:/WT/GC/206.pdf&Open=True> para 2.6.

17 Compulsory licensing is legislated for in the UK under sections 48, 48A and 48B of the Patents Act 1977 (as amended).

provisions – services of the state (in Ireland) and for the crown (under UK law)[18] – which are similar in nature. Hence, the analysis of the domestic framework applicable in Ireland for compulsory licensing and, particularly, licensing for services of the state resonates with the framework applicable in the UK context.[19]

The article is structured as follows: part one examines the potential impacts patents have on access to healthcare, providing a case for why effective compulsory licensing measures are needed at the national level to alleviate access issues posed by patents in the health context. Part two then examines the overarching international framework for compulsory licensing applicable in all WTO states including Ireland. This WTO framework sets down minimum criteria which present restrictions on uses of compulsory licences at a national level for COVID-19 which Ireland must continue to abide by in any reform of national laws in this area or face the possibility of WTO dispute settlement proceedings and potential trade sanctions. Following this, part three considers the domestic compulsory licensing framework applicable in Ireland, offering a critique of how the current framework would apply to the COVID-19 context, including the shortcomings evident. It then offers proposals for how this framework could be reformed to facilitate a more effective compulsory licensing system whilst remaining compliant with Ireland's international obligations. Part four examines licensing for service of the state provisions under Irish law, focusing on their potential to be used in the COVID-19 context, and reforms which would make this system more effective. Part five then outlines practical obstacles for the use of compulsory licences in Ireland posed by EU laws around data/marketing exclusivity and the EU's opt-out of the relevant WTO framework under Article 31*bis*. This section argues that such issues need to be addressed to ensure they do not cause undue barriers to compulsory licensing within EU states. Finally, part six concludes by arguing that compulsory licensing and government-use provisions are important tools within a state's broader arsenal of devices to alleviate access issues posed by patents on health-related technologies, including within the COVID-19 context. Accordingly, it is vital that national laws facilitate effective systems for such licensing interventions. The Irish state should address existing shortcomings as soon as possible lest such licensing measures be required for COVID-19.

1 Patents, access to health and COVID-19: The need for effective national compulsory licensing mechanisms

Patents and patent-holder decisions on licensing can have a significant impact on access to health-technologies.[20] Once a technology is patented the patent-holder has the discretion to dictate how that invention is used and by whom for the duration of the patent (generally 20 years).[21] The patent-holder can refuse to license a technology to third-parties, which could effectively mean the patent-holder becomes the sole provider of that technology and, depending on their manufacturing capacity, this can have knock-on implications for the supply of that technology. This in turn can have significant adverse implications for health if the underlying patented technology is a medicine, vaccine, or diagnostic and if supply of such technologies is limited. These issues have

18 See section 55, 56 and 59 Patents Act 1977, as amended in the UK context.

19 It is, however, acknowledged that the two jurisdictions differ on some aspects, and this must be borne in mind in drawing any comparative lessons from the analysis for the UK context.

20 For a general discussion of the potential impacts of patents on access to and delivery of healthcare, which argues in favour of greater oversight of patent-holder's discretion in this context, see McMahon (n 4).

21 TRIPS Agreement, Article 33 which states that: 'The term of protection available shall not end before the expiration of a period of twenty years counted from the filing date.'

been brought into the spotlight by COVID-19, particularly in light of recent debates around vaccine/medicine nationalism,[22] where some states have sought to negotiate agreements for preferential access of proposed vaccines or treatments for COVID-19 in their national states,[23] prioritising national interests over the interests/needs of other states. Such deals can have knock-on effects causing shortages of supplies of such vaccines, medicines or diagnostics available in other states.[24] Moreover, these types of deals are only likely to increase as we get closer to finding effective vaccines/treatments for COVID-19, with particularly acute effects on poorer nations and vulnerable populations.[25]

To alleviate these issues, states can try to *encourage* patent-holders to share intellectual property rights via voluntary licensing initiatives as part of a broader global solidarity approach to tackle COVID-19. Notable examples of such initiatives within the COVID-19 context include the Open COVID-19 pledge, and the World Health Organization's (WHO) COVID Technology Access Pool (CTAP).[26] It is undoubtedly important for states to endorse such initiatives,[27] and this article is not arguing that compulsory licensing offers a substitute to replace voluntary licensing initiatives. Rather, it argues that such voluntary mechanisms to encourage sharing of intellectual property should be strongly supported and endorsed by states, as state support for such voluntary licensing initiatives can form an important nudge to encourage patent-holders to participate in such initiatives, and the more national states do so, the greater the pressure placed on patent-holders to commit to voluntary licensing initiatives for COVID-19.

However, voluntary licensing models do not replace the need to have effective compulsory licensing mechanisms because voluntary licensing initiatives by their nature are subject to patent-holder opt-in, and, whilst many patent-holders may be willing to engage with these initiatives for COVID-19, they are not generally legally mandated to do so, and some patent-holders inevitably will not. Similarly, whilst many patent-holders may of their own accord be agreeable to license their COVID-19 technology on favourable

22 See discussion in McMahon (n 6).

23 Stephen Buranyi, '"Vaccine nationalism" stands in the way of an end to the Covid-19 crisis' *The Guardian* (London, 14 August 2020) <www.theguardian.com/commentisfree/2020/aug/14/vaccine-nationalism-stands-in-the-way-of-an-end-to-the-covid-19-crisis>; Donato Paolo Mancini and Michael Peel , '"Vaccine nationalism" delays WHO's struggling Covax scheme' *Financial Times* (London, 2 September 2020) <www.ft.com/content/502df709-25ac-48f6-aee1-aec7ac03c759>.

24 Sarah Boseley, 'US secures world stock of key Covid-19 drug remdesivir' *The Guardian* (London, 30 June 2020); Barbara Mintzes and Ellen 't Hoen, 'The US has bought most of the world's remdesivir. Here's what it means for the rest of us' (*The Conversation*, 3 July 2020) <https://theconversation.com/the-us-has-bought-most-of-the-worlds-remdesivir-heres-what-it-means-for-the-rest-of-us-141791>. Similar attempts to acquire preferential supplies at a national level are evident around proposed COVID-19 vaccines. See Duncan Matthews, 'Coronavirus: how countries aim to get the vaccine first by cutting opaque supply deal' (*The Conversation*, 27 July 2020) <https://theconversation.com/coronavirus-how-countries-aim-to-get-thevaccine-first-by-cutting-opaque-supply-deals-143366>.

25 See also discussion on potential impacts on vulnerable populations in Ana Santos Rutschman, 'How "vaccine nationalism" could block vulnerable populations' access to COVID-19 vaccines' (*The Conversation*, 17 June 2020) <https://theconversation.com/how-vaccine-nationalism-could-block-vulnerable-populations-access-to-covid-19-vaccines-140689>.

26 See <https://opencovidpledge.org>. For information on CTAP, see: <www.who.int/emergencies/diseases/novel-coronavirus-2019/global-research-on-novel-coronavirus-2019-ncov/covid-19-technology-access-pool>. An overview of both initiatives is provided in McMahon (n 6). See discussion of distinction between pools and pledges in: J L Contreras, M Eisen, A Ganz et al, 'Pledging intellectual property for COVID-19' (2020) 38 Nature Biotechnology 1146, 1147.

27 The benefits of voluntary licensing initiatives in such contexts, and the need for state support of these are discussed in detail in McMahon (n 6).

terms, others may not. Accordingly, it is vital that states have an effective avenue to allow third parties or governments to use patented health-technologies without patent-holder consent where this is needed to alleviate access issues for COVID-19. Compulsory licensing and licensing for service of the state mechanisms provide such avenues, allowing states to intervene with a patent-holder's control over patented health-technologies where patent-holders refuse to license or provide such technologies on reasonable terms.[28] Accordingly, where state support for voluntary licensing initiatives perform a useful 'carrot' function encouraging patent-holders to share intellectual property rights on reasonable terms, compulsory licensing is a vital 'stick' that states can use as a threat or as an mandatory measure where patent-holders do not offer reasonable terms and where public health requires access to patented technologies.[29] In effect, arguably, voluntary and compulsory licensing measures are complementary in nature and serve different functions in the access to medicines context.

Put simply, compulsory licensing measures are useful in such contexts because if a patent-holder refuses to license the patented technology on reasonable terms, depending on the national patent laws applicable, a compulsory licence could be applied for in that state which, if granted, would allow the government/third-party to produce that health-technology for supply within that state.[30] Alternatively, provisions allowing licences for service of the state where applicable (such as within Ireland) could be used to authorise the government to produce such technologies for that state. Such mechanisms, as noted, are useful negotiation tools, as the threat by the state of issuing a compulsory licence if reasonable licensing terms cannot be reached can be sufficient to encourage patent-holders to adopt more favourable licensing terms.[31] Hence, such licensing mechanisms provide a useful avenue to alleviate access issues where public health needs require these,[32] resetting the balance of control over how a patented technology is licensed away from patent-holders to states where needed. Accordingly, it is vital that existing national mechanisms for licensing without patent-holder authorisation are examined and reformed as necessary to ensure they are as effective as possible to meet public health needs.

It is acknowledged that compulsory licensing measures are conducted at the national level, and, where such licences are granted, they are predominantly aimed at alleviating issues in the national state. Thus, some may question whether using such mechanisms is akin to *vaccine/medicine nationalism* by another means. However, two points can be made to this potential critique: 1) Compulsory licences issued at the national level can give rise to a momentum for change which has global benefits, causing patent-holders to voluntarily license patented technologies for use elsewhere or to commit to not enforcing their

28 Such mechanisms could also be used in cases of vaccine/treatment nationalism to alleviate shortages of medicines which may arise elsewhere if patent-holders provide preferential supplies of patented vaccines, treatments, or diagnostics for particular states. See Mintzes and 't Hoen (n 24).

29 For a discussion of such measures as 'carrots and sticks' and the broader corporate responsibility context of patents, see Aisling McMahon and Edana Richardson, 'Patents, health and corporate responsibility' (Working Paper 2020) (on file with the author).

30 The issues may be more complex in the vaccine context as, alongside compulsory licences, other information on the process may be needed to allow a third party sufficient knowledge to produce the vaccine. Nonetheless, compulsory licences can help alleviate at least part of the issue posed in such contexts.

31 See discussion in Gorik Ooms and Johanna Hanefeld, 'Threat of compulsory licences could increase access to essential medicines' (2019) 365 British Medical Journal l2098.

32 See also discussion of role and use of such compulsory licensing flexibilities in Ellen 't Hoen, Jacquelyn D Veraldi, Brigit Toebes and Hans Hogerzeil, 'Medicine procurement and the use of flexibilities in the Agreement on Trade-Related Aspects of Intellectual Property Rights 2001–2016' (2018) 96 Bulletin of the World Health Organization 185–193.

intellectual property rights – we have already seen this in the COVID-19 context for the drug Kaletra, discussed below. 2) Arguably, the greater the number of states that use compulsory licences, the more normalised compulsory licensing becomes within the patent system for public health emergency contexts such as pandemics. This can make it easier for other states to use such mechanisms, and, relatedly, the greater number of states that threaten to use compulsory licensing within a particular context, the higher the likelihood of reputational damage to patent-holders who refuse to voluntarily license their patents on reasonable terms, which again can act as a strategy to encourage effective global change. Therefore, although such compulsory measures take place at the national level, they can have much broader, global or regional benefits.

Importantly, in making such arguments, the article is not suggesting that compulsory licensing or licensing for service of the state mechanisms are a panacea to address all access issues posed by intellectual property rights within the COVID-19 context. It is conceded that patents are not the only form of intellectual property rights relevant in the COVID-19 healthcare context, and particularly within the vaccine space, where trade-secret protection is also important.[33] In many cases, how a vaccine is produced may be protected by trade-secret protection and this knowledge may be necessary to create a generic version of a vaccine. This in turn differentiates vaccines from small-molecule medicines which are often easier to replicate by others without additional knowledge, e.g. of the manufacturing process, from the patent-holder.[34] In the vaccine context or for biological (complex) medicines, having a compulsory licence over the patent will not necessarily *on its own* enable a third party to produce a similar version of that vaccine/medicine without the third party also having access to the information protected under the trade secret or additional know-how. This information may not be disclosed by the patent-holder, thereby requiring the third party to develop this knowledge to produce a generic vaccine – and this may be difficult and/or take considerable time. Nonetheless, this is not a reason to dismiss the role of compulsory licensing, as, even if it is more difficult to replicate a vaccine/medicine in such contexts, removing the patent obstacle will bring a third party closer to doing so. Furthermore, although outside the scope of this article, such issues arguably merely support the argument that, once a compulsory licence is issued on a patented invention, patent-holders should also disclose related information around the working of that invention, such as additional know-how and trade secret information.[35]

It is also acknowledged that alongside ensuring *effective* measures for compulsory licensing, there is a broader issue around states' willingness to use such compulsory

33 Trade secrets are defined by the World Intellectual Property Office as 'intellectual property rights on confidential information which may be sold or licensed'. See <www.wipo.int/tradesecrets/en/>. See also discussion in David S Levine, 'Covid-19 should spark a re-examination of trade secrets' stranglehold on information' (*Stat News*, 10 July 2020) <www.statnews.com/2020/07/10/covid-19-reexamine-trade-secrets-information-stranglehold>.

34 Sara Eve Crager, 'Improving global access to new vaccines: intellectual property, technology transfer, and regulatory pathways' (2014) 104 American Journal of Public Health S414–S420. For a discussion on the difference between small molecule and biologic medicines and why manufacturing process information is information, see Generics and Biosimilars Initiative <http://www.gabionline.net/Biosimilars/Research/Small-molecule-versus-biological-drugs>.

35 See discussion of possible avenues to obtain access to trade secret protected information under TRIPS in David S Levine, 'Covid-19 trade secrets and information access: an overview' (*Infojustice*, 10 July 2020) <http://infojustice.org/archives/42493>. See also discussion in McMahon (n 6).

licensing measures in the COVID-19 context and for public health more generally. Historically, higher-income states have been reluctant to use such compulsory licensing measures given the potential threat of backlash within the international community. Indeed, countries such as Thailand, Brazil and India faced backlash from the US and elsewhere, including threats of trade sanctions, for using such measures.[36] Some states, such as Ireland, with strong pharmaceutical industries may be particularly reluctant to use compulsory licensing measures, fearing industry backlash. However, it is in all our interests to eradicate the virus as soon as possible, from both a health and an economic industry perspective, and this can only be achieved by ensuring effective access to COVID-19 health-technologies.

Furthermore, arguably, COVID-19 is acting as a catalyst for change in this context, as a trend towards greater acceptability of the use, and need for, effective licensing interventions in the pandemic context is evident. For example, as noted, many countries have already amended laws to ensure effective avenues to obtain compulsory licences are available where needed for COVID-19.[37] Moreover, Israel issued the first compulsory licence for the pandemic on 18 March 2020, to allow it to import generic versions of AbbVie's Kaletra for COVID-19,[38] rather than receiving backlash/criticism, soon after Israel issued this compulsory licence, AbbVie committed to not enforcing its patents over Kaletra globally for COVID-19.[39] Given this backdrop, states, including Ireland, should be encouraged to evaluate existing systems and, arguably, will be more willing and empowered to introduce reforms and use compulsory licensing where needed for COVID-19.

Thus, in short, whilst compulsory licensing measures are not a panacea to solve all issues around access to health-technologies posed by intellectual property rights for COVID-19, they are a vital tool to alleviate access issues posed by patents. It is therefore imperative that states ensure such compulsory mechanisms are as effective as possible at a national level, so that they are open to states to use where needed to address one part of the broader access puzzle.

2 Licensing without patent-holder's authorisation: Ireland's international obligations

Turning then to the international framework within which licences without patent-holder's authorisation (such as compulsory licences and licences for service of the state) can be granted. At an international level, the minimum criteria for the grant of such licences are set out under Article 31 of the TRIPS Agreement. The TRIPS Agreement is applicable in all 164 WTO states globally, including Ireland. The EU is also a signatory of

---

36 See C T Scopel and G C Chaves, 'Initiatives to challenge patent barriers and their relationship with the price of medicines procured by the Brazilian Unified National Health System' (2016) 32 Cad Saude Publica 121; S Tantivess, N Kessomboon and C Laongbua, 'Introducing government use of patents on essential medicines in Thailand, 2006–2007: policy analysis with key lessons learned and recommendations' (International Health Policy Program 2008); Z Siddiqiu, 'India defends right to issue drug "compulsory licenses"' (*Reuters*, 23 March 2016), as cited in E 't Hoen et al, 'Medicine procurement and the use of flexibilities in the Agreement on Trade-Related Aspects of Intellectual Property Rights, 2001–2016' (2018) 95 Bulletin of the World Health Organization 185, 189.

37 Houldsworth (n 11); see discussion in Hilary Wong, 'The case for compulsory licensing during COVID-19' (2020) 10(1) Journal of Global Health 010358 <doi:10.7189/jogh.10.010358>.

38 Houldsworth (n 11).

39 Phil Taylor, 'AbbVie won't enforce patents for COVID-19 drug candidate Kaletra' (*PharmaForum*, 25 March 2020) <https://pharmaphorum.com/news/abbvie-wont-enforce-patents-for-covid-19-drug-candidate-kaletra>.

the TRIPS Agreement, which makes the TRIPS Agreement 'binding upon the institutions of the Union and its Member States'.[40] Moreover, whilst the TRIPS Agreement does not have direct effect in EU member states by virtue of EU law,[41] nonetheless, when domestic courts are interpreting provisions of the TRIPS Agreement, they must do so 'as far as possible' 'in the light of the wording and the purpose' of the TRIPS provision, given the EU entered into the TRIPS Agreement on behalf of states.[42] Alongside the provisions within the TRIPS Agreement, all contracting parties of the TRIPS Agreement must also comply with the substantive provisions of the Paris Convention for the Protection of Industrial Property 1883, as amended (hereafter the Paris Convention) of which Article 5 relates to licensing without the rights-holder's authorisation.[43] Thus, the TRIPS Agreement and Paris Convention set out minimum standards for the use of compulsory licences within WTO states. Moreover, because the EU is a signatory of the TRIPS Agreement, there is an additional legal impetus arising from EU law for EU member states to abide by the provisions within TRIPS.

However, patent law is jurisdictional in nature and there is no global patent system *per se*. Instead, in practice, a compulsory licence is obtained at the national level, and the processes to apply for compulsory licences are governed by national laws separately in each state. Therefore, whilst every WTO member state must abide by relevant provisions within the TRIPS Agreement and the Paris Convention which set down *minimum* standards, they are free to impose *higher* standards than required under these international treaties within national laws for mechanisms for licensing without patent-holder's authorisation. Differences can therefore arise in the applicable rules around compulsory licensing in each national context. Accordingly, it is only by considering the requirements under the Paris Convention, the TRIPS Agreement and under national laws in the state concerned that a picture of how such licensing mechanisms operate in each national context, such as in Ireland, can be gleaned.

This section considers the international framework applicable for the grant of licences without patent-holder authorisation, focusing specifically on the minimum criteria set out under the Paris Convention, and under the TRIPS Agreement. It identifies the main shortcomings with this international law framework in terms of how such international laws might impact the use of compulsory licensing mechanisms in Contracting States such as Ireland within the COVID-19 context. However, such restrictions are in many cases presented as offering a balance between patent-holders' right to intellectual property and broader public interests concerns.[44] Furthermore, changes to this international framework are likely to be a longer-term project given the complexity of achieving change

40 Article 216(2) Treaty on the Functioning of the EU, as discussed in Justine Pila and Paul Torremans, *European Intellectual Property Law* (Oxford University Press 2017) 69.

41 Pila and Torremans (n 40) 69.

42 Joined Cases C-300/98 and 392/92 *Parfums Christian Dior SA v TUK Consultancy BV* [2000] I-11307 [47] as discussed ibid.

43 Many WTO states, including Ireland, were already party to the Paris Convention upon the adoption of the TRIPS Agreement in 1995. The Paris Convention came into force in Ireland on 4 December 1925.

44 On the right to intellectual property within the European Convention on Human Rights system, see Christophe Geiger and Elena Izyumenko, 'Intellectual property before the European Court of Human Rights' in Christophe Geiger, Craig A Nard and Xavier Seuba (eds), *Intellectual Property and the Judiciary* EIPIN series vol 4 (Edward Elgar 2018) 9–90. One could question the extent to which this balance incorporates other human rights concerns, such as the right to health. See, generally, discussion in Laurence R Helfer, 'Human rights and intellectual property: conflict or coexistence?' (2003) 5 Minnesota Intellectual Property Review 47; Philippe Cullet, 'Human rights and intellectual property protection in the TRIPS Era' (2007) 29(2) Human Rights Quarterly 403–430.

of the TRIPS framework.[45] Within the COVID-19 context, therefore, it is more fruitful for the Irish government in the shorter term to ensure the legal framework in Ireland offers the most effective national framework for compulsory licensing possible, whilst remaining compliant with the current TRIPS framework. This national framework is discussed in part three below.

2.1 Compulsory licensing and the Paris Convention 1883

The Paris Convention was signed in 1883, however, the original Convention did not refer to compulsory licensing of patents. Compulsory licensing was discussed in the 1925 Revision Conference and subsequently provisions were included in the Paris Convention setting out minimum requirements for the grant of such licences.[46] In its current version, Article 5 of the Paris Convention states that contracting states of the Convention: '*shall* have the right to take legislative measures providing for the grant of compulsory licences to prevent the abuses which might result from the exercise of the exclusive rights conferred by the patent, for example, failure to work'.[47] This provision marks a recognition that contracting states are permitted to grant compulsory licences under the Convention. The Convention does not provide an exhaustive list of grounds under which a compulsory licence can be granted, merely providing 'failure to work' an invention as one example of a ground that a compulsory licence could be granted for. However, it is permissible under the Paris Convention for states to adopt other grounds for compulsory licences beyond failure to work.[48] This provides considerable discretion to states in relation to the grounds which can be adopted at a national level for compulsory licensing.

Notably, the Paris Convention provides that a compulsory licence cannot be applied for:

> on the ground of failure to work or insufficient working *before the expiration of a period of four years from the date of filing of the patent application or three years from the date of the grant of the patent*, whichever period expires last; it shall be refused if the patentee justifies his inaction by legitimate reasons …[49] (emphasis added)

This provision means that national states are restricted in providing a compulsory licence *where it is granted on the ground of 'failure to work or insufficient working'* until a period of three years after grant or four years from the date of application has passed. However, this time restriction is not applicable beyond the circumstance of where a licence is granted for failure to work/insufficient working, i.e. it does not apply if compulsory licences are granted on other grounds. This point is returned to below in discussing the national requirements applicable under Irish law for the grant of a compulsory licence.

2.2 Compulsory licensing and the TRIPS Agreement

Alongside these provisions within the Paris Convention, there are seven main cumulative requirements for a compulsory licence under the TRIPS Agreement relevant to the healthcare context and to COVID-19.[50] Each of these criteria are considered here, alongside the implications of these criteria for the use of compulsory licensing for patented health-technologies for COVID-19.

---

45 Frederick M Abbott, 'The future of the multilateral trading system in the context of TRIPS' (1997) 20 Hastings International and Comparative Law Review 661, 667.

46 Esther Van Zimmeren and Geertrui Van Overwalle, 'A paper tiger? Compulsory license regimes for public health in Europe' (2011) 42 International Review of Intellectual Property and Competition Law 1.

47 Article 5(A)(2) Paris Convention 1883, as amended.

48 Van Zimmeren and Van Overwalle (n 46).

49 Article 5(A)(4) Paris Convention 1883, as amended.

50 For a full list of requirements: Article 31, TRIPS Agreement.

Firstly, under the TRIPS Agreement, each authorisation of a compulsory licence must be considered 'on its individual merits'.[51] This implies that a state, for instance, cannot issue a blanket compulsory licence for an area of technology or specific issue such as issuing a compulsory licence for all 'COVID-19 related medicines'. Instead, each application for use of a patent without the patent-holder's permission must be considered individually for each individual medicine, vaccine etc. This requirement rules out the ability of any WTO state, including Ireland, using compulsory licences in a blanket manner to facilitate access to COVID-19 health-technologies.

Secondly, a compulsory licence can only be granted if the proposed user of the licence had previously tried to obtain an authorisation for use of that technology from the patent-holder on reasonable commercial terms and conditions, and such efforts were unsuccessful within a 'reasonable period of time'.[52] A state can waive this requirement in 'the case of a national emergency or other circumstances of extreme urgency or in cases of public non-commercial use'. However, the rights-holder must be notified as soon as reasonably practicable in the context of a national emergency or situation of extreme urgency. Moreover, in the context of authorisation for public non-commercial use, the rights-holder 'shall be informed promptly'.[53]

The COVID-19 context could likely fall within the definition of a national emergency, particularly if a state of emergency was declared by the country where the compulsory licence was sought (although an official declaration of a state of emergency is not necessarily required for this waiver to apply).[54] Arguably, the fact that COVID-19 was declared a global pandemic by the WHO in March 2020,[55] and as this pandemic continues this may be sufficient to constitute a national emergency for the purposes of Article 31(b) TRIPS Agreement. This is supported by the fact that the text of the Doha Declaration on the TRIPS Agreement and Public Health, para 5(c) states that:[56]

> Each member has the right to determine what constitutes a national emergency or other circumstances of extreme urgency, it being understood that public health crises, including those relating to HIV/AIDS, tuberculosis, malaria and other *epidemics*, can represent a national emergency or other circumstances of extreme urgency [emphasis added].

Moreover, if there were a shortage of supplies of patented COVID-19 medicines, vaccines or diagnostics, this would arguably constitute a circumstance of extreme urgency given the likely threat to life such a shortage would pose within that state. In such circumstances, as in the case of a country seeking to provide COVID-19 medicines, vaccines or diagnostics to the public on a non-commercial basis, the requirement of having previously attempted to negotiate a licence with the patent-holder could be waived.

---

51 Article 31(a) TRIPS Agreement.

52 Article 31(b) TRIPS Agreement.

53 Ibid.

54 For discussion of states of emergency in the COVID-19 context, see Alan Greene, 'State of emergency: how different countries are invoking extra powers to stop the coronavirus' (*The Conversation*, 30 March 2020) <https://theconversation.com/state-of-emergency-how-different-countries-are-invoking-extra-powers-to-stop-the-coronavirus-134495>.

55 'WHO Director-General's opening remarks at the media briefing on COVID-19' (11 March 2020) <www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>.

56 See WTO, Ministerial Declaration of 14 November 2001, WTO Doc WT/MIN(01)/DEC/1, 41 ILM 746 (2002) (hereinafter Doha Declaration).

Third, the scope and the duration of the licence must be limited to the purpose for which it was authorised.[57] This implies that, after a crisis has been averted – or in the case of a compulsory licence granted in the COVID-19 context, after the global pandemic has been deemed to have ceased – then the compulsory licence would need to terminated (or within a reasonable period after this). This could be problematic within the health context because if a compulsory licence ceases it would likely reintroduce access issues, transferring control back to the patent-holder and thereby affecting who can supply such health-technologies thereafter. It could also lead to an increase in costs and limits on the supply of such patented medicines, vaccines or diagnostics. This could have particularly significant implications for developing countries that have less access to resources/funds to secure access to such products.

Fourth, the authorisation or licence must be non-exclusive, meaning that the compulsory licence does not stop the patent-holder from licensing others to use the technology.[58]

Fifth, any use under a compulsory licence is authorised 'predominantly for the supply of the domestic market' of the state where that use is authorised in.[59] This requirement will effect countries that have limited manufacturing capacity to make patented vaccines, medicines or diagnostics domestically even if a compulsory licence were granted. The implications of this requirement were previously evident in the AIDS crisis during the 1990s, and as a result the Doha Declaration was adopted,[60] and, subsequently, Article 31*bis* was introduced, which allows states to import patented inventions made under a compulsory licence elsewhere under certain circumstances.[61] However, shortcomings remain in relation to such provisions which have been well documented elsewhere.[62] Moreover, the EU has opted out of this procedure which means an EU state cannot be an eligible importing member, which could prove highly problematic in the COVID-19 context – this is a serious and unnecessary shortcoming for compulsory licensing use within EU states discussed in detail in part five below.

Sixth, under the TRIPS Agreement the patent-holder must be paid adequate remuneration in each case.[63] This provision may deter states/third parties from issuing compulsory licences, as there is uncertainty around how 'adequate remuneration' should be determined within states.[64] To facilitate better use of compulsory licensing, where public health demands it, further guidelines on how such remuneration should be

57 Article 31 (c) TRIPS Agreement.

58 Article 31(d), TRIPS Agreement.

59 Article 31(f), TRIPS Agreement.

60 Paragraph 6 of which stated: '6. We recognize that WTO members with insufficient or no manufacturing capacities in the pharmaceutical sector could face difficulties in making effective use of compulsory licensing under the TRIPS Agreement. We instruct the Council for TRIPS to find an expeditious solution to this problem and to report to the General Council before the end of 2002.' Doha Declaration (n 56).

61 Article 31 Bis, TRIPS Agreement, as amended. See Doha Declaration (n 56).

62 For a discussion, see Ellen 't Hoen, 'TRIPS, pharmaceutical patents and access to essential medicines: Seattle, Doha and beyond' (2002) 3(1) Chicago Journal of International Law <https://chicagounbound.uchicago.edu/cjil/vol3/iss1/6>; Duncan Matthews, 'WTO decision on implementation of paragraph 6 of the DOHA Declaration on the TRIPs Agreement and public health: a solution to the access to essential medicines problem?' (2004) 7 Journal of International Economic Law 73.

63 Article 31(h), TRIPS Agreement.

64 See discussion in Antony Taubman, 'Rethinking TRIPS: "adequate remuneration" for non-voluntary patent licensing' (2008) 11(4) Journal of International Economic Law 927–970.

determined would be useful,[65] alongside ensuring clarity on this at a national level – a point returned to below.

Seventh, the compulsory licence is liable to termination 'if and when the circumstances which led to it cease to exist and are unlikely to recur'.[66] Moreover, the 'competent authority shall have the authority to review, upon motivated request, the continued existence of these circumstances'.[67] Thus, the grant of such a licence can and indeed should be subject to periodic review to ascertain if the circumstances for grant still pertain. This implies that compulsory licensing provisions are time-limited in nature, and, for example, if COVID-19 was no longer considered a global pandemic or a national emergency (if this was the basis on which such a licence was granted), then that licence shall be liable to termination. Similarly, if a licence were granted in circumstances of urgency, once these circumstances dissipated then the licence would be liable to termination.

From the foregoing, it is thus evident that these culminative minimum WTO criteria present constraints on the use of compulsory licensing which may impede their effectiveness in alleviating access issues posed by patents for COVID-19.[68] Most notably, from the above, the provisions around the need for adequate remuneration for the rightsholder could deter applicants/states in using compulsory licensing, unless greater guidance is evident on how this would be determined. Furthermore, the need to grant a compulsory licence on a case-by-case basis for each patented invention means compulsory licensing does not offer a catch-all solution for addressing access issues on, for example, health-technologies for COVID-19. Nonetheless, taking a pragmatic view, it is likely to be difficult to successfully petition for any change of such TRIPS provisions, given that this would require a change of WTO law, which is fraught with difficulty given the need for multiple states' agreement, and, even if successful, any such change would likely take years to achieve. Furthermore, as noted, such conditions offer a balance to ensure interference with rights-holder's patent rights is not disproportionate. Thus, changing WTO law is not likely a feasible solution within the COVID-19 pandemic context, although addressing some of these shortcomings should not be abandoned in the longer term as doing so would facilitate states to offer a more effective system for compulsory licensing at the national level in future where public health requires such interventions.

Nonetheless, whilst WTO states including Ireland must abide by these international TRIPS and Paris Convention standards, national frameworks require consideration to ensure that compulsory licensing provisions are not going beyond the requirements set out within the TRIPS and Paris Convention, if the aim is to ensure compulsory licences offer effective mechanisms where needed to facilitate access to patented health-technologies. Changes to national laws are feasible within a shorter period of time, and, indeed, since COVID-19 many national states have already adopted legal measures to amend national laws on compulsory licensing to ensure such licences are easier to obtain

---

65 WHO guidelines on this and proposals for how remuneration could be calculated are available at: James Love, *Remuneration Guidelines for Non-voluntary Use of a Patent on Medical Technologies* WHO/TCM/2005.1 (WHO 2005) <https://apps.who.int/iris/handle/10665/69199>

66 Article 31(g) TRIPS Agreement.

67 Article 31(g) TRIPS Agreement

68 See also Frederick M Abbott and Rudolf Van Puymbroeck, 'Compulsory licensing for public health: a guide and model documents for implementation of the Doha Declaration paragraph 6 decision' (2005) World Bank Working Paper 61 <http://documents.worldbank.org/curated/en/173701468337882214/Compulsory-licensing-for-public-health-a-guide-and-model-documents-for-implementation-of-the-Doha-Declaration-Paragraph-6-Decision>.

if required in the COVID-19 context.[69] Ireland should follow suit, removing national obstacles to effective uses of compulsory licensing for COVID-19 and for public health more generally.

3 Compulsory licensing in Ireland and COVID-19: An appraisal and proposal for reform

This section offers an overview of the current applicable framework for compulsory licensing in Ireland, and the main shortcomings of this framework, focusing specifically on how compulsory licensing under the current framework could be used to alleviate access issues posed by patents for COVID-19 health-technologies. It then sets out reform proposals to offer a more effective system for compulsory licensing in Ireland.

Prior to delving into the analysis of such provisions, it is important to note two points. Firstly, compulsory licensing provisions operate in Ireland alongside a government-use provision, allowing state use of a patented invention where needed for 'services of the State' (examined in part four below). These two avenues should be viewed as a package of complementary measures which operate together to permit interventions with patent-holder's control over patented inventions where needed to alleviate access issues arising.

Secondly, on first view, considered together, the texts of these legal provisions on compulsory licensing and licensing for service of the state in Ireland, although as will be seen they do not refer to public health or the public interest explicitly, nonetheless appear to offer scope for intervention with patent-holder discretion, particularly, because many of the provisions related to compulsory licensing and licences for 'service of the State' are drafted in a broad 'open-textured' manner.[70] Thus, one could argue, that because of the relatively expansive nature of some of these provisions (and particularly the service of the state provisions), they *could in theory* allow a broad interpretation, including for such provisions in their current form to be used, in some circumstances, to alleviate access issues in a pandemic context.

However, two issues arise in this context which support the need for reform. Firstly, Irish law contains instances where higher standards for compulsory licensing have been adopted than are needed to comply with minimum TRIPS standards. As will be seen, this limits the scope for use of such licences than could be possible in the COVID-19 context, and reforms are needed to address this.

Secondly, given the historical reluctance of states to issue compulsory licences, such expansive open-textured legislative provisions as they exist in the Irish context which fail to expressly refer to the use of these provisions in the health or pandemic context, and which require the Controller of Intellectual Property in Ireland or courts to interpret the measures as applying in the pandemic context, are problematic. This is because 'thought styles',[71] or interpretative communities,[72] whereby shared understandings or predispositions in favour of particular actions/interpretations may arise within such interpretative contexts. Within such contexts, the past actions within a decision-making framework can affect future decision-making, and such influences can be particularly

69 Houldsworth (n 11).

70 Open-textured provisions are defined by Hart as 'statutes may be a mere legal shell and demand by their express terms to be filled out with the aid of moral principles'. See H L A Hart, *The Concept of Law* (Oxford University Press 1961) 199–200. See discussion in Aisling McMahon, 'Morality provisions in the European patent system: an institutional examination' (PhD Thesis, University of Edinburgh, 2016).

71 Mary Douglas, *How Institutions Think* (Syracuse University Press 1986).

72 Peter Drahos, 'Biotechnology patents, markets and morality' (1999) 21(9) European Intellectual Property Review 441–442.

acute where decision-makers have considerable discretion over open-textured or expansive provisions.[73] In the institutional context, this is discussed in terms of 'path dependency' which implies that historical actions influence present acts,[74] or in simple terms that 'what happened at an earlier point in time will affect the possible outcomes of a sequence of events occurring at a later point in time'.[75] In essence, this implies that how a particular issue or provision (or an analogous issue) has been dealt with in the past by the institution will be influential, but not necessarily determinative, of present action(s).[76]

Accordingly, in situations where there has been a historical reluctance to interpret open-textured provisions in a manner which limits patent rights,[77] or where there has been historically limited use of such provisions to intervene with patent-holder discretion over licensing, this *status quo* of non-intervention and hence non-use/grant of compulsory licences is likely to be maintained unless change is encouraged by external action, such as by legislative reforms which specify that such provisions can be used in public health or pandemic contexts. In Ireland, there have been no compulsory licences issued to date in such contexts under section 70 of the Patents Act 1992 (as amended) (hereafter PA), and, similarly, the service of the state provision has received limited attention or use. Hence, leaving such provisions in their current form, without legislative intervention or guidance that expressly indicates the use of such provisions is possible within the public health and/or in a pandemic or epidemic context, is arguably likely to result in the *status quo* remaining and entrench a lack of willingness to use or interpret these provisions as applying in the public health context in Ireland.

For these reasons and for the specific reasons given below, legislative change is urgently needed to clarify the operation of these provisions in the public health context and to address specific shortcomings within the current framework.

3.1 Compulsory licensing in Ireland

Section 70 of the PA[78] sets out the main criteria for the grant of a compulsory licence in Ireland. It provides that any person can apply for a compulsory licence to the Controller of Intellectual Property in Ireland,[79] (hereafter Controller) based on two main grounds namely:

1. on the basis of demand in the Irish state:
    a. 'a demand in the State for the subject matter of the patent is not being met or is not being met on reasonable terms';[80] or

73 See generally McMahon (n 70).

74 See O Hathaway, 'Path dependence in the law: the course and pattern of legal change in a common law system' (2001) 86 Iowa Law Review 101.

75 W H Sewell Jr, 'Three temporalities: toward an eventful sociology' in T McDonald, (ed), *The Historic Turn in the Human Sciences* (University of Michigan Press 1996) 245, 262–263.

76 See McMahon (n 70); Aisling McMahon, 'Regulatory authorities and decision-making in health research: the institutional dimension' in Graeme Laurie et al (eds), *Cambridge Handbook of Health Research Regulation* (Cambridge University Press 2021).

77 See McMahon (n 70).

78 This was amended by the Patent (Amendment) Act 2006; see also Patent Rules 1992, as amended, of which rule 50 applies in the context of compulsory licences.

79 The name of the Controller was changed from 'Controller of Patents Designs and Trade Marks' to the 'Controller of Intellectual Property' from December 2019, and this was amended in the Patents Act 1992, by virtue of section 42 of the Copyright and Other Intellectual Property Law Provisions Act 2019.

80 Section 70(1)(a)(i) PA.

b. demand in the Irish state for the patented product is 'being met by importation other than from a member of the World Trade Organization';[81]

2. that the 'establishment or development of commercial or industrial activities in the State is unfairly prejudiced'.[82]

Of these grounds, the first ground is the ground most likely applicable in the COVID-19 context. However, the fact that grant of a compulsory licence in Ireland under this ground is premised on the demand for the invention in the state could pose an obstacle to the use of a compulsory licence, as assessing 'demand in a State' is a relatively subjective assessment. Under previous case-law, 'demand' in this context has been understood to mean public demand.[83] However, this would likely be construed as the actual demand at the price provided by the patent-holder not demand if the invention was provided at a lower price.[84] In the pandemic context, demand is likely to be high even if a vaccine, medicine or diagnostic is charged at high prices, given that access to this health-technology is required to meet a pressing health need. Thus, the need to show actual demand is not likely to be an obstacle. However, it is unclear what threshold is needed to show that demand is not being met under this ground. Arguably, unless a significant shortage of the patented invention were evident it may be difficult to justify a compulsory licence on this ground. Nonetheless, the lack of clarity around what is likely to constitute a failure to meet demand or failure to meet demand on 'reasonable terms' (or whether this could plausibly include a consideration of the price the invention is provided for) could deter applications or authorisations for compulsory licences in Ireland.

Alongside these two main grounds, a compulsory licence can also be obtained if a patent-protected invention (the second patent) cannot be used in a state without infringing upon rights from another patent (the first patent) – in such cases the patent-holder of the second patent can apply to the Controller for a licence under the 'first patent to the extent necessary for the exploitation of the invention concerned, provided that such invention involves an important technical advance of considerable economic significance in relation to the invention claimed in the first patent'.[85]

Overall, these grounds for compulsory licensing in Ireland are relatively restrictive in nature, even though, as noted above, the Paris Convention and the TRIPS Agreement provide considerable discretion to states in setting the grounds for compulsory licences. The restrictive nature of these grounds forms the first potential roadblock within Irish laws to providing an effective mechanism for compulsory licensing in the health context, and this is returned to in section 3.2 below.

Relatedly, section 70 PA, as amended, specifies that compulsory licences can only be granted three years after the publication of notice of the grant of a patent in Ireland.[86] This mirrors the criteria for the grant of compulsory licences in cases of 'failure to work or insufficient working' of a patent under the Paris Convention, discussed above. This condition

81 Section 70(1)(a)(ii) PA.

82 Section 70 (1)(b) PA.

83 *Boult's Patent* (1909) 26 RPC 383 as cited in Clark et al (n 12) 201.

84 Ibid 201.

85 Section 70(2) PA.

86 Section 70(1) PA states: 'At any time after the expiration of the period of three years, or such other period as may be prescribed, beginning on the date of the publication of notice of grant of a patent any person may apply to the Controller for a licence under the patent, or for an entry in the register to the effect that licences under the patent are to be available as of right …'.

of time within the Paris Convention allows the patent-holder sufficient time to show they are exploiting the patent and hence are not failing to work an invention.[87] However, in the Irish context, the reference to this time restriction covers all grounds for compulsory licensing and is not restricted to the context of 'failure to work' a patented invention, the context where this time-limit is specified as applying under the Paris Convention. This is an example of an additional restriction within Irish law beyond those which are required by the international legal framework applicable. To be compliant with international laws, it would be sufficient for Irish law to include this time restriction *only* in relation to a ground of grant of a compulsory licence based on failure to work an invention.[88]

This time restriction is potentially an obstacle to the use of compulsory licensing for COVID-19 related health-technologies, as it limits the state's ability to grant a compulsory licence for newly developed patented technology in Ireland. Generally, within the health-context, the regulatory approval process for new medicines or vaccines takes considerable time, often well beyond three years.[89] Thus, ordinarily having to wait until three years after patent grant to apply for a compulsory licence over a health-technology may not in practical terms pose a significant obstacle to accessing this technology. However, time is of the essence for COVID-19. There is a global effort to develop effective and safe medicines and vaccines for COVID-19 as soon as possible, yet this three-year requirement means that, even if there is such a medicine or vaccine, if it is under patent no compulsory licence can be applied for until three years post-grant. This obstacle should be removed as soon as possible outside the circumstances of where it is required under the Paris Convention (i.e. failure to work a patent). As having this time restriction for all grounds of compulsory licensing in Ireland unnecessarily ties the hands of the government for newly patented medicines/vaccines/diagnostics.

Turning then to the terms upon which the grant of a compulsory licence is made in Ireland, if the Controller is satisfied the grounds for a compulsory licence are met, according to section 70(3) PA the licence can be granted with the following main conditions:

- it would be a non-exclusive licence;
- it would be granted predominantly for the supply of the domestic market in Ireland;
- the licence granted can only be assigned with prior authorisation of the Controller and under specific conditions;
- the licence is granted subject to the payment to the patent-holder of 'adequate remuneration in the circumstances of the case, taking into account the economic value of the licence'. In this context, it is questionable how the adequacy of remuneration is to be determined, including how the 'economic value' of the licence is to be calculated, which could lead to uncertainty and potential challenge. Providing guidelines on this within Ireland would be useful to ensure this is not a roadblock to use of compulsory licensing if needed for COVID-19;

---

87 Van Zimmeren and Van Overwalle (n 46) 17.

88 There was such a specific ground under Irish patent law, under section 70(2)(a) of the Patents Act 1992 which stated that: '(2) The grounds referred to in subsection (1) are the following: (a) that the invention which is the subject of the patent, being capable of being commercially worked in the state, is not being commercially worked therein or is not being so worked to the fullest extent that is reasonably practicable.' This was deleted by section 19, Patents Amendment Act 2006.

89 Stuart A Thompson, 'How long will a vaccine really take?' *New York Times* (New York, 30 April 2020) <www.nytimes.com/interactive/2020/04/30/opinion/coronavirus-covid-vaccine.html>.

- the licence granted would be limited in terms of the scope and purpose for which it is granted.[90]

Many of these requirements merely reiterate the minimum requirements set out under the TRIPS Agreement considered above and are necessary to ensure Ireland's compliance with such international obligations.

In granting a compulsory licence under the Act, the Controller shall take account of the following matters:

a. the nature of the relevant invention, the time which has elapsed since the grant of the patent and the measures already taken by the proprietor or any licensee to make full use of the invention,

b. the ability of any person to whom a licence would be granted under the order to exploit the patent to the public advantage, and

c. the risks to be undertaken by that person in providing capital and exploiting the patent if the application is granted.[91]

These aspects imply a focus on the supply of the invention within the country by the patent-holder and suggest a need to allow the patent-holder time to develop sufficient supply of the invention. This condition likely relates to the narrow grounds currently in operation for compulsory licensing in Ireland. However, significantly, this section does not reference any public need or the public interest more generally, and, arguably, including a reference to demand to meet public need/interest would be preferable should compulsory licences be required for public health purposes.

Section 72 PA expressly provides that a Minister of the Irish government can apply for a compulsory licence under any of these above grounds set out under section 70 'after the expiration of the period of three years beginning on the date of the publication of notice of grant of a patent, or such other period as may be prescribed under section 70(1)' where section 70(1) provides that it can also be after 'such other period as may be prescribed'. Presumably, this caveat in section 70(1) allows the government to adopt legislative measures which would adopt an alternative period which could be used to reduce the time needed in circumstances outside of cases where compulsory licences are being used to address failure to work or failure to work the patent on sufficient basis.

Under section 72 PA, the minister can apply for the entry in the patent register to the effect that:

> … licences under the patent are available as of right, or for the grant to any person specified in the application of a licence under the patent, and the Controller may, if satisfied that any of those grounds are established, make an order in accordance with the application.[92]

Thus, section 72 PA allows, for example, a health minister to apply for a compulsory licence for medicines/vaccines or other technology necessary in the COVID-19 context.

In terms of the practical application process for a compulsory licence, applicants apply to the Controller and provide a statement which sets out the facts of the application including 'evidence indicating that the applicant sought to obtain a licence' from the patent-holder but was unable to obtain such a licence on 'reasonable terms and within a reasonable time'.[93] The Controller '*may*, when so requested' (emphasis added) by the

90 Section 70(3) PA.

91 Section 70(4) PA.

92 Section 72(1) PA.

93 Section 73(1) PA.

applicant dispense of this requirement for evidence if 'there exists a national emergency or other circumstances of extreme urgency' or 'in the case of an application for a licence for public non-commercial use' provided that the patent-holder has been 'informed as soon as reasonably practicable of the intention of the applicant to apply to the Controller for licence under patent'.[94] This mirrors provisions allowing for waiver of the requirement of attempting to negotiate access with the patent-holder in cases of emergency, urgency or public non-commercial use as set out in TRIPS.

3.2 Reform proposals

In short, based on the foregoing analysis, to facilitate greater effectiveness of the compulsory licensing mechanisms within Ireland for the COVID-19 context and for public health more generally, it is proposed that four main aspects of the current national framework be reconsidered, namely two legislative changes and two areas where further guidance on how legislative provisions are interpreted in practice would be useful

In the context of legislative reforms: firstly, it is vital that the requirement to wait for three years after the patent grant to apply for a compulsory licence is reconsidered. This requirement is only necessary under the Paris Convention in cases where a patent is applied for based on failure of the patent-holder to work the patent, or insufficient working of the patent.[95] However, the time restriction has been applied for all grounds of compulsory licensing applicable in Ireland under section 70 PA, thereby going beyond what is needed for Ireland to comply with international law, and this should be amended. This could be achieved by amending section 70(1) of the Act. Alternatively, currently, section 70(1) PA provides that it is a three-year term 'or such other period as may be prescribed', implying this could be amended relatively easily by prescribing a shorter term by virtue of, for example, a statutory instrument. For the reasons given above, reducing this term is needed to ensure the Irish government's hands are not tied in using compulsory licensing for newly patented inventions. This time restriction should therefore be removed under Irish law for all grounds of compulsory licensing, other than in cases where the licence is granted on the basis of failure to work a patent or failure to work it sufficiently within Ireland. Thus, alongside amending section 70(1) to remove the time restriction, to ensure compliance with the Paris Convention, the Act could reintroduce a ground which expressly referred to the grant of a compulsory licence in the case of failure to work a patent or insufficient working of a patent[96] and should, in that context, include a provision which stated: 'in cases where a compulsory licence is granted on the basis of failure to work or insufficient working of a patent, such a licence shall only be granted 3 years after publication of notice of patent grant'.

Secondly, and relatedly, the grounds allowing for compulsory licences in Ireland could be expanded to include a broader catch-all ground relevant to the public interest/public health context. For example, a ground could be included which allowed for a compulsory licence where 'necessary for the public interest'. A non-exhaustive list should also be given within such an amendment, providing that it would apply, 'for example, in cases of

94 Section 73(1A) PA.

95 Article 5(A)(4) Paris Convention.

96 As noted above, a ground which expressly referred to failure to work an invention was evident in the Patents Act 1992 but was removed by the Patents Amendment Act 2006 in favour of a more generally worded ground related to demand in the state. Section 70(2)(a) Patents Act 1992 previously provided the following ground for a compulsory licence: 2(a) 'that the invention which is the subject of the patent, being capable of being commercially worked in the State, is not being commercially worked therein or is not being so worked to the fullest extent that is reasonably practicable'.

national emergency, including, within a public health crisis (e.g. a pandemic), environmental emergency, or economic crisis'. As noted above, expressly referring to such measures as applicable in the public health context, including in a pandemic context, within the legislative framework would provide an explicit source referring to the use of such measures in the public health context, which could encourage greater willingness to use provisions where needed for COVID-19. Expanding the text to include examples such as use in an environmental emergency, economic crisis etc. would provide greater longevity for the proposed reform expressly confirming its use, where public interest required it, in other contexts. However, adopting a broad provision of this kind may be opposed by industry, and an alternative would be to adopt a provision which allowed compulsory licensing where necessary for public health grounds, with a legislative provision specifying that an epidemic or pandemic context constitutes an example of public health grounds.[97] Or, alternatively, given that compulsory licensing and service of the state provisions operate as a package in Ireland, one solution would be to refer expressly to the public health context only under patents for 'service of the State' and not in the general compulsory licensing context. However, such an approach would unnecessarily confine the grounds within which licences could be granted for public health purposes, requiring service of the state provisions to be fulfilled, and this would fail to facilitate the use of compulsory licensing more broadly on this basis. It would also fail to facilitate third parties applying for compulsory licences on public health grounds, as licences for service of the state relate primarily to government use of such provisions.

Turning then to the need for further guidance: thirdly, certain aspects of the current grounds would benefit from further guidance, specifically under section 70(a) PA, guidance is needed on how 'demand' for the invention in the state will be assessed in practice, including what threshold is needed. Greater clarity is also needed around how whether an invention is being provided on 'reasonable' terms will be considered in practice, and how 'reasonableness' will be assessed.

Fourthly, and finally, guidelines should be issued within Ireland on how the requirement of adequate remuneration for a compulsory licence will be assessed in a particular context, ensuring any uncertainty in this context does not act as a deterrent to applicants in applying for compulsory licences or the approval of such licences. Whilst remuneration must be decided on a case-by-case basis in practice, having a statement on the overarching general principles applicable within Ireland on this, and raising awareness of the mechanisms used to determine this or general principles applicable, would arguably be useful to encourage greater recourse to such provisions, where needed.

The need for such legislative reforms and guidance around existing provisions under Irish law is compounded by the fact that the grant or applications for compulsory licences have historically been uncommon/non-existent in practice.[98] Thus, there is limited jurisprudence on compulsory licensing to draw on to give further clarity in the Irish context. Moreover, patent legislation in Ireland has been amended several times since the Patent Act's adoption in 1992, including a relatively substantial amendment of the compulsory licensing provisions in 2006 to bring Ireland's laws in line with the TRIPS Agreement. Thus, any case law which does exist will likely be of limited use in determining the application of recently revised provisions. Moreover, whilst, the law in the UK is likely to be persuasive, given its similar wording to much of the current Irish

97 Other countries have adopted such provisions in light of COVID-19. See, generally, Wong (n 37).

98 There are no official statistics on this in Ireland: see European Patent Academy, *Compulsory Licensing in Europe: A Country-by-Country Overview* (European Patent Office, 2018) 63.

patent legislation,[99] nonetheless, to introduce greater clarity in the area, further domestic guidance would be beneficial.

4 Licensing for service of the state in Ireland: Access to health-technologies and COVID-19

Alongside the compulsory licensing provisions applicable under Irish law, section 77 PA provides for the right to use patented inventions in Ireland without the patent-holder's permission for 'service of the State' – this is effectively a government-use provision (akin to the crown-use within UK law).[100] This provision allows a government minister or person authorised by them powers to do any of the following acts without the consent of the patent-holder:

(a) where the invention is a product, make, use, import or stock the product or dispose of or sell or offer to dispose of or sell it to any person;

(b) where the invention is a process, use it or do in relation to any product obtained directly by means of the process anything mentioned in paragraph (a);

(c) supply or offer to supply to any person any of the means, relating to an essential element of that invention, for putting the invention into effect.[101]

This mechanism provides a relatively broad avenue within which use of a patented technology can be sought by the government without the patent-holder's consent and this could be useful within the COVID-19 context. Any of the above acts if conducted under this provision do not amount to patent infringement.[102]

The use of the invention under this section for 'service of the State' is defined as 'a service financed out of moneys charged on or advanced out of the Central Fund or moneys provided by the Oireachtas or by a local authority for the purposes of the Local Government Act, 1941'.[103] Arguably, this would undoubtedly include Irish health services, given the public nature of such services in Ireland as provided by the Health Services Executive (HSE). In terms of renumeration, the TRIPS Agreement provides that a state can set out specific circumstances where it can use a patented invention without the patent-holder's consent, provided 'adequate' remuneration is paid to the rightsholder.[104] The PA provides that use of the invention under this section is subject to terms which may be agreed upon 'either before or after the use' with the approval of the Minister for Finance, by any government Minister and the applicant for the patent or the patent-holder.[105] Presumably, this also includes agreement on the terms of remuneration which could therefore be concluded before or after use under this section. In default of an agreement, or in the event of dispute, the matter would be settled under the Act by

99 See discussion ibid.

100 For a discussion of the operation of crown-use in the context of COVID-19, see Karen Walsh, Andrea Wallace, Mathilde Pavis, Natalie Olszowy, James Griffin and Naomi Hawkins, 'Intellectual property rights and access in crisis' (Working Paper 2020) (on file with author). Other countries have similar crown or government-use provisions, for a list of such measures and their use in the COVID-19 context to date, see COVID-19 IP Policy Tracker <www.wipo.int/covid19-policy-tracker/?utm_source=WIPO+Newsletters&utm_campaign=d5be8aff4c-EMAIL_CAMPAIGN_2020_05_28_07_40&utm_medium=email&utm_term=0_bcb3de19b4-d5be8aff4c-256647357#/covid19-policy-tracker/access>. For a general discussion of these provisions under Irish law, see: Clark et al (n 12) [8.46]–[8.53] .

101 Section 77(1) PA.

102 Section 77(2) PA.

103 Section 77(10) PA.

104 Article 31(h) TRIPS Agreement, as amended. See discussion in Clark et al (n 12) 8.48.

105 Section 77(3) PA.

the appropriate court or by an arbitrator upon conditions they may direct.[106] However, a shortcoming within this process is if the terms of remuneration for the patent-holder were not concluded until after the use takes place, and, if there is uncertainty around the principles applicable for how remuneration is calculated, the government would not know how much it would likely cost to use this licensing mechanism, and this could potentially lead to a government's reluctance to use this provision, as high costs could arise.[107] Similar, to the compulsory licensing context, such issues would be alleviated to some extent by guidelines being provided at a national level around the overarching principles applicable for how remuneration would likely be calculated or factors which would be considered within this process in Ireland. This could act as a useful toolkit in deciding adequate remuneration and would introduce greater clarity on these issues, thereby making it more likely that a government would use these provisions if needed.

Section 78(1) PA provides for additional extended provisions for uses of inventions for service of the state in exceptional circumstances, whereby the state can use patented inventions for any purposes where it appears necessary or expedient to a minister of the government for: 'a) maintenance of supplies and services essential to the life of the community; b) for securing a sufficiency of supplies and services essential to the well-being of the community … f) for ensuring the public safety and the preservation of the State'. Although, the text of these provisions does not expressly refer to public health or the public interest, it is highly likely in theory that such provisions would be applicable within the COVID-19 context given the global pandemic we face, as it is likely that supplies of diagnostics, treatments or vaccines for COVID-19 would be viewed as supplies which were 'essential to the life of the community' and/or 'essential to the well-being of the community' and/or for 'the preservation of public safety in the State'. Thus, such grounds could potentially offer an effective avenue to gain access to such patented technology where necessary for COVID-19.[108]

However, the thresholds applicable under these circumstances would again benefit from further guidance. In particular, the use of these provisions (a) and (b) when applied to the COVID-19 context would likely rest on whether adequate supplies of a patented health-technology were available in the country. Yet, it is questionable how the notion of 'sufficiency of supplies' in this context would be interpreted. Greater legislative guidance on the threshold applicable would be useful in the event that this provision is needed for COVID-19 or within other health contexts in future.[109]

A further issue is that section 78(2) indicates the powers to use a patented invention for service of the state under the section 78(1) provision would only be invoked in

106 Section 77(6) PA. Use of an invention for service of the state can be used without renumeration under the circumstances of section 77(4) PA which provides: '(4) Where an invention which is the subject of any patent or application for a patent has, before the date of filing, or, where priority is claimed, the priority date of the application, been duly recorded in a document by, or been tried by or on behalf of any Minister of the Government (such invention not having been communicated directly or indirectly by the applicant for or the proprietor of the relevant patent), any Minister of the Government or such of his officers, servants or agents as may be authorized in writing by him, may use the invention so recorded or tried for the service of the State free of any royalty or other payment to the applicant for or the proprietor of the patent, notwithstanding the existence of the application or patent ...'

107 For a discussion of the analogous Crown use provision in the UK context, see Walsh et al (n 100).

108 Notably, in *Evalve & Abbott v Edwards Lifesciences Limited* [2020] EWHC 513 (Pat), [77] where the similar UK mechanism of crown-use was raised: the court noted that an example of where this provision would be applicable was for the provision of life-saving generic medicines in the public interest in 'special cases, such as novel pandemic disease'. See discussion in Walsh et al (n 100).

109 See discussion in Kelly (n 21).

exceptional circumstances. It states that: 'Where the Government are of the opinion that, owing to the *existence of exceptional circumstances*, it is desirable in the *interests of the community* that a power conferred by subsection 1 shall be available, they may by order declare that the power shall be available.' (emphasis added)[110] This current phrasing explicitly suggests use under section 78(1) grounds will only be in rare circumstances,[111] which may prove a considerable impediment to use of this mechanism as it suggests a very high threshold is needed if it is only to be used in 'exceptional circumstances'. It is proposed that legislative change is needed and the phrase the 'existence of exceptional circumstances' in this section be amended to indicate 'Where the Government are of the opinion that, it is *necessary in the interests of the community* that a power conferred by subsection 1 shall be available, they may by order declare that the power shall be available.' This should be accompanied within the legislation by a non-exhaustive list of examples of what i*nterests of the community* may include, with express reference to public health contexts, including a public health emergency such as a pandemic/epidemic etc.

Without adopting such changes, it is still likely that the global pandemic caused by COVID-19 would fall within exceptional circumstances suggested within section 78 PA, but expressly providing within legislation that a pandemic or epidemic comprises a circumstance where such licences if needed would fall within the interests of the community would arguably encourage greater clarity and, hopefully, result in greater state willingness to use such provisions where needed. Removing the term 'exceptional circumstances' would also potentially encourage greater recourse to this provision. Nonetheless, the elephant in the room may be the remuneration required if a licence were issued under this provision – and guidance on how this is determined is therefore vital.

5 EU obstacles to effective use of compulsory licensing

Alongside these national reform proposals, it is also vital that some of the current obstacles stemming from EU law to the effective practical use of compulsory licensing in the health context within EU states are amended.[112] Most notable in this context are existing legal protections which pose difficulties for the registration of generic medicines (medicines which are similar/identical to the branded/patented version) that are produced under compulsory licences, namely data exclusivity protections and marketing exclusivity protections. Under EU law, there is an eight-year data exclusivity protection which applies to all new medicines[113] – this effectively means that within this time period

110 Section 78(2(a) PA.

111 See also discussion in Ann Henry, 'Coronavirus: patents rights and the public interest' (Pinsent Masons, Out-Law, 3 April 2020) <www.pinsentmasons.com/out-law/analysis/coronavirus-patents-rights-public-interest>.

112 Ellen 't Hoen, 'European pharmaceutical legislation needs exceptions to data and market exclusivity to protect european patients from high drug prices' (*Medicines Law and Policy*, 21 May, 2018) <https://medicineslawandpolicy.org/2018/05/european-pharmaceutical-legislation-needs-exceptions-to-data-and-market-exclusivity-to-protect-european-patients-from-high-drug-prices>.

113 See Regulation (EC) No 726/2004 of the European Parliament and of the Council of 31 March 2004 laying down Community procedures for the authorisation and supervision of medicinal products for human and veterinary use and establishing a European Medicines Agency OJ L 136, 30.4.2004, Article 14(11) which states: 'Without prejudice to the law on the protection of industrial and commercial property, medicinal products for human use which have been authorised in accordance with the provisions of this Regulation shall benefit from an eight-year period of data protection and a ten-year period of marketing protection, in which connection the latter period shall be extended to a maximum of 11 years if, during the first eight years of those ten years, the marketing authorisation holder obtains an authorisation for one or more new therapeutic indications which, during the scientific evaluation prior to their authorisation, are held to bring a significant clinical benefit in comparison with existing therapies.'

someone who, for example, produces a generic medicine under a compulsory licence would be unable to use the original clinical data conducted for the patented medicine's approval to support the application for generic approval. Having to conduct additional clinical studies would be both costly and time prohibitive for a generic producer and could effectively defeat the purpose of granting a compulsory licence as, in practice, although it could mean the licence to allow a third party to produce a generic version is granted, it would be difficult to gain approval for generics made under the compulsory licence. There is currently no explicit waiver on the data exclusivity protection where a medicine is subject to a compulsory licence.[114] This is a major regulatory stumbling block to the use of compulsory licensing in the EU for COVID-19 and other health contexts. Furthermore, marketing exclusivity applies, which means that a generic medicine cannot be marketed until 10 years after the original medicine obtained authorisation.[115] There is also no exception to this marketing exclusivity protection under EU law in cases of compulsory licensing. Such data and marketing exclusivity protections deterred Romania's use of a compulsory licence in 2016 for sofosbuvir to treat hepatitis C.[116] In order for compulsory licensing measures to be an effective avenue in EU countries for health-technologies, it is crucial that a waiver is introduced to such protections in this context.

Finally, there are obstacles to Ireland to importing medicines manufactured under compulsory licence in another country where needed for COVID-19. As noted above, Article 31 of the TRIPS Agreement allows for compulsory licensing, but it requires that such licences are used *predominantly for the supply of the domestic market* within the state that the licence is granted within.[117] This can be highly problematic as, for example, in the developing-country context states may not have the domestic manufacturing capacity to produce generic versions of medicines under compulsory licence in the state. To address this, a wavier to this domestic production requirement for compulsory licensing was introduced by Article 31*bis*, which allows states to produce generic versions of health-technologies made under compulsory licence to be exported to states that require them but do not have the manufacturing capacity to produce these domestically, subject to certain requirements. This waiver was first introduced on a temporary basis on 30 August 2003; it was then approved as an amendment to the TRIPS Agreement in 2005, and finally ratified by the required number of WTO states, to enter into force in January 2017. This waiver is the only avenue by which states can import generic versions of patented medicines from another state where public health demands it, whilst remaining compliant with the TRIPS Agreement.

This provision is coming under the spotlight again in the COVID-19 context, for two reasons: 1) as has been seen already, COVID-19 can cause devasting health impacts within a region, and this in turn could impact domestic industries including pharmaceutical production. Thus, states that ordinarily have a strong pharmaceutical industry could find this halted by COVID-19 or other public health contexts, making it impossible to produce sufficient generics domestically even if a compulsory licence were granted. 2) States may

114 See discussion in Houldsworth (n 11). See also, generally, discussion of data exclusivity and compulsory licensing in Phoebe Li and Pheh Hoon Lim, 'A precautionary approach to compulsory licencing of medicines: tempering data exclusivity as an obstacle to access' (2014) 3 Intellectual Property Quarterly 241.

115 Article 14(11) Regulation 726/2004.

116 E 't Hoen, P Boulet, B Baker, 'Data exclusivity exceptions and compulsory licensing to promote generic medicines in the European Union: a proposal for greater coherence in European pharmaceutical legislation' (2017) 10 Journal of Pharmaceutical Policy and Practice 19.

117 Article 31(f) states that: 'any such use shall be authorized predominantly for the supply of the domestic market of the Member authorizing such use'.

lack domestic supplies of pharmaceutical ingredients necessary to make generic medicines for COVID-19, again rendering domestic manufacture of COVID-19 health-technologies under compulsory licence impossible.

Importantly, although historically this provision was most relevant to developing countries and least developed countries that lacked manufacturing capacity, it can be used by any WTO state that has notified the WTO that it is an eligible importing state, whereby 'it being understood that a Member may notify at any time that it will use the system in whole or in a limited way, for example only in the case of a national emergency or other circumstances of extreme urgency or in cases of public non-commercial use'.[118] However, the Annex to this provision also states that '*some Members will not use the system as importing Members*, and that some other Members have stated that, if they use the system, it would be in no more than situations of national emergency or other circumstances of extreme urgency' (emphasis added).

The EU is one such region that has voluntarily opted-out of the WTO framework, and therefore EU states are not eligible importing states, meaning the EU or member states cannot currently import generics made under compulsory licence elsewhere under Article 31*bis*.[119] Ireland has a strong pharmaceutical industry, and thus, ordinarily, this opt-out decision would be unlikely to pose an issue as there should be sufficient capacity/skills to produce generic medicines made under compulsory licence domestically. However, in the context of a global pandemic which can have devastating impacts on human health,[120] with further impacts on supply chains,[121] this could conceivably hinder any country's ability to continue pharmaceutical operations. Therefore, despite Ireland's ordinarily strong pharmaceutical manufacturing capabilities, it is highly plausible Ireland could need to utilise this measure for COVID-19 or future health crises. Thus, it is vital this opt-out decision for EU states is changed, as James Love has noted in this context, it is:

> … totally irrational for any country, even a rich country, to keep its own hands tied to meet the COVID-19 needs of its population by voluntarily shutting itself off from patented ingredients, components, and essential medical products and supplies.[122]

There are calls for the EU to change this restriction, and the Irish government and other national EU Member State governments should support such calls to demand change on this as a matter of urgency.[123] The opt-out should either be revoked in its entirety if

118 Annex to the TRIPS Agreement, paragraph 1(b).

119 General Council Decision, Article 1(b), footnote 3; Article 31*bis*, Annex, Article 1(b), footnote 3 <www.wto.org/english/tratop_e/trips_e/implem_para6_e.htm>.

120 At the time of writing, 4 September 2020, COVID-19 had resulted in over 873,000 deaths worldwide, with 26.5 million cases. See <www.worldometers.info/coronavirus>

121 See discussion in Anthony Lakavage, 'Covid-19 has exposed cracks in the global medicines supply chain. We need to fix them' (*Stat News*, 2 June 2020) <www.statnews.com/2020/06/02/covid-19-exposed-cracks-global-medicines-supply-chain>.

122 James Love, 'Open letter asking 37 WTO Members to declare themselves eligible to import medicines manufactured under compulsory licence in another country, under 31bis of TRIPS Agreement' (Knowledge Ecology International, 7 April 2020) <www.keionline.org/32707> accessed 28 August 2020.

123 Ibid; see discussion in: Christopher Garrison, 'Never say never – why the high income countries that opted-out from the Art 31bis WTO TRIPS system must urgently reconsider their decision in the face of the COVID-19 pandemic' (*Medicines, Law and Policy*, 8 April 2020) <https://medicineslawandpolicy.org/2020/04/never-say-never-why-the-high-income-countries-that-opted-out-from-the-art-31bis-wto-trips-system-must-urgently-reconsider-their-decision-in-the-face-of-the-covid-19-pandemic/>.

possible, or the EU should ensure a provision is included within this opt-out which allows such imports in cases of 'national emergency or other circumstances of extreme urgency',[124] so that importation of health-technologies made under compulsory licence elsewhere can be used where needed by EU states within the COVID-19 context and other health emergency situations.

Conclusion

Given the devastating global impacts of the COVID-19 pandemic on human life, society and the broader economy, it is vital that affordable global access to vaccines, diagnostics and treatments for COVID-19 are available once these are developed, as such access must be delivered without delay. However, as has been demonstrated, patents, depending on how patent-holders choose to use them, can prove significant obstacles to such access. Accordingly, it is vital that national governments have effective mechanisms in place to temper patent-holder control via compulsory licensing and government-use licences allowing states to intervene over patent-holder decision-making where needed. Such measures are also effective as negotiation tools for states, as often the threat of such licences will encourage patent-holders to offer access to patented health-technologies on more favourable terms.

Accordingly, the Irish government must re-evaluate the current framework for compulsory licensing and licensing for service of the state in Ireland to ensure the national legal framework applicable is as effective as possible, whilst remaining compliant with Ireland's international obligations. As demonstrated, shortcomings exist within the current framework in this context, and these should be addressed as soon as possible to offer a broader compulsory licensing system in Ireland which gives the state greater scope to intervene with patent-holder's licensing decisions. More specifically, the following changes are needed:

1. legislative change to amend the grounds for compulsory licensing offered under section 70 PA;
2. legislative change to remove the three-year time restriction outside of the context of compulsory licensing for insufficient working of a patent;
3. offering further guidance at a national level on what is meant within the current grounds in relation to 'demand' in the state; and
4. offering national guidance and raising awareness of how 'adequate' remuneration will be determined with reference to overarching principles applicable.

The article also proposes that the current provisions around licensing for service of the state are amended to incorporate:

1. legislative guidance making it clearer what threshold of supply or lack of supply is needed under section 78(a) and (b) PA;
2. legislative change removing the reference to 'exceptional circumstances' under section 78(2) and an express acknowledgment within the legislation that a pandemic/epidemic may constitute circumstances where licences for services of the state can be granted in the 'interests of the community' under this section; and
3. further guidance at a national level on the overarching principles used to determine adequate remuneration in such contexts.

124 Ibid.

Finally, obstacles to effective uses of such compulsory licensing mechanisms posed by EU data exclusivity and marketing exclusivity protections should be addressed as a matter of urgency, as these present significant practical impediments to using such licensing avenues for COVID-19 and for other public health crises. Ireland should petition the EU for a waiver to be provided within such protections where compulsory licences or licences for service of the state are issued. Ireland should also join calls for the removal of the EU's opt-out of the WTO framework allowing states to import medicines produced elsewhere under compulsory licensing.

In short, it is in all our interests, both in the health and economic sense, that COVID-19 is eradicated as soon as possible. It is also in all our interests that any obstacles to this eradication are addressed, and this includes obstacles created by patent rights. Compulsory licensing and licences for service of the state are an important tool to alleviate such access issues, which if left unchecked may significantly impede future access to supplies of medicines, vaccines and improved diagnostics for COVID-19. Allowing such potential patent obstructions to subsist without an effective avenue for remedy is contrary to all our interests, hence, using such licensing measures and adopting legislative provisions and national guidance to make the system for their use more effective within Ireland and other states is not only appropriate but wholly warranted.

Abortion activism, legal change, and taking feminist law work seriously

MÁIRÉAD ENRIGHT

KATHRYN MCNEILLY

FIONA DE LONDRAS*

Abstract

Abortion laws in the Republic of Ireland and Northern Ireland have recently undergone radical reform. This occurred following a 2018 referendum in the Republic and the passing of the Northern Ireland (Executive Formation etc) Act 2019 in Northern Ireland. In both jurisdictions, these legal changes are the products not only of moments of constitutional and legislative action or of litigation, but of decades of feminist protest and strategising that both generated and exploited moments of legal opportunity. In this article, drawing on a 2018 workshop and qualitative interviews with feminist activists, we focus attention on what we call the 'feminist law work' involved in reform, highlighting the role of non-lawyer activists in achieving legal change in instrumental, creative, emotional, and laborious ways. We argue that 'feminist law work' should be taken seriously as a highly skilled and indispensable driving force in formal legal change processes.

Keywords: abortion; activism; law reform; legal change; feminist law work; feminist legal studies; Ireland; Northern Ireland.

> **Interviewer:** If I had to ask you, right now, what's going on – what are you doing to change the law [on abortion], what would your answer be?
>
> **Kellie O'Dowd (activist, Alliance for Choice):** What are we not doing?

Introduction

Over the last two years, the law on access to abortion has changed substantially in both jurisdictions on the island of Ireland. In the Republic of Ireland, the 8th Amendment to the Constitution was repealed and replaced with a constitutional provision permitting the Oireachtas to regulate the termination of pregnancy, which was achieved with the Health (Regulation of Termination of Pregnancy) Act 2018. In Northern Ireland, abortion was

* Reader in Feminist Legal Studies, Birmingham Law School; Senior Lecturer, Queen's University of Belfast School of Law; Professor of Global Legal Studies, Birmingham Law School.

This research was supported by research activity grants from Birmingham Law School and Ulster University School of Law, a research activity grant from the SLSA to support interview transcription (Enright), and the Philip Leverhulme Prize (de Londras, PLP-2017-181). We are grateful to Dr Catherine O'Rourke of the Transitional Justice Institute, University of Ulster, for co-organising the workshop and securing funding to support it, as well as for her discussions on the animating ideas of the paper. Our thanks also to all of our interviewees, and to everyone who attended the workshop, for their time and their feminist law work.

decriminalised in 2019,[1] and is now legally available pursuant to new regulations introduced in 2020.[2] In both cases, the changes were brought about by particular legal 'events' – constitutional referendum, legislative change, the passage of regulations – but those events were preceded by decades of feminist activism and agitation,[3] including substantial amounts of what we term 'feminist law work'.[4] This is work that takes place alongside, and sometimes against, more formalised law work and activity undertaken by voluntary and non-governmental organisations. It is apposite, now, to reflect on the dynamics of legal change in abortion law reform in Ireland and, particularly, to attend to and recognise the role that feminist law work has played in bringing that change about.

As explained further below, the term 'feminist law work' is intended to capture the manifold ways in which non-lawyer feminist activists engage with law, legal mobilisation, legal consciousness, legal institutions and legal argumentation in order to try to achieve their desired end of safe, legal, and local abortion for pregnant people who seek it. Our aim in this paper is not to provide a comprehensive account of the decades of feminist law work that was undertaken in Ireland or, indeed, to suggest that formal legal changes have brought that feminist law work to an end. Instead, using structured conversations with eight invited abortion rights activists from Ireland and Northern Ireland as our jumping-off point, we seek to contribute to the complex retelling of the stories of abortion law reform in Ireland and Northern Ireland.

The paper proceeds in five main parts. First, we outline our ethical and methodological approach (Part 1), before briefly describing the prevailing legal and political context in which the feminist law work that our participants discuss took place (Part 2). We then look in some detail at three different elements of feminist law work that emerged especially clearly in our engagement with activists and which shaped our understanding of the phenomenon: ambivalence about law (Part 3), the multifaceted nature of feminist law work (Part 4), and feminist law work's 'everydayness' (Part 5).

1 Our ethical and methodological approach

In this article we investigate the feminist law work undertaken by eight non-lawyer activists whom we invited to a workshop held in Ulster University, Belfast, on 23 November 2018. The participants were engaged in different ways in abortion rights activism, with three working primarily in the Republic of Ireland and five in Northern Ireland. Our workshop took place at a crucial turning point: Irish abortion law had been liberalised and there was a strong sense (as the slogan went at the time) that 'the North is Next'.[5] Northern Irish feminist activists were at this point concentrating their attention on lobbying for abortion law reform at Westminster. Before the workshop, and in preparation for the discussion to be held there, Enright interviewed a number of the participants. These pre-workshop interviews helped us to shape the workshop discussion, which was oriented towards identifying how the activists to whom we spoke thought

1 Northern Ireland (Executive Formation etc) Act 2019.

2 Abortion (Northern Ireland) Regulations 2020.

3 Fiona Bloomer, 'Protests, parades and marches: activism and extending abortion legislation to Northern Ireland' in Lisa Fitzpatrick (ed), *Performing Feminisms in Contemporary Ireland* (Carysfort Press 2013) 245–266; Ruth Fletcher, 'Silences: Irish women and abortion' (1995) 50(1) Feminist Review 44.

4 See e.g. Sally Sheldon, Jane O'Neill, Clare Parker and Gayle Davis, '"Too much, too indigestible, too fast?" The decades of struggle for abortion law reform in Northern Ireland' (2020) 83(4) Modern Law Review 761.

5 Melissa Davey, '"North is next": fresh fight for grassroots power that beat Ireland abortion ban' *The Guardian* (London, 1 June 2018) <https://www.theguardian.com/world/2018/jun/02/north-is-next-fresh-fight-for-grassroots-power-that-beat-ireland-abortion-ban>.

about and enacted their relationship with law. The concept of feminist law work emerged from the workshop, rather than driving it. As we discussed what these women had spoken about, how they thought about law, and how they had trained themselves to challenge, manipulate, shape, undermine, ridicule and reform law as part of their activism, it became clear to us that what they had described was a vital form of legal labour rooted in feminist commitments and playing itself out in their everyday engagements with law, politics, reproduction, (il)legality, and activist strategising. Ethical approval for this qualitative research was granted by the University of Birmingham. All workshop participants and interviewees gave full informed consent to their participation in the research as well as use of the resulting data in published outputs. All participants reviewed a draft of this article prior to final publication.

The participants in our workshop and interviews were women who have been active in struggles to reform abortion law in Northern Ireland and the Republic of Ireland, and in the spaces between them. Ailbhe Smyth has been involved in feminist and reproductive rights activism for more than 40 years, is head of the Coalition to Repeal the 8th Amendment, and was one of the co-directors of Together for Yes, the official civil society platform for a 'Yes' vote in the May 2018 referendum. Linda Kavanagh is a long-time leader within the Abortion Rights Campaign and worked in social media for Together for Yes. Erin D'Arcy is an artist and founder of 'In Her Shoes'; an intervention in the referendum campaign which curated women's anonymous first-hand stories of being denied abortion care in Ireland. She published these stories on Facebook and, near the end of the campaign, in a hard copy form for distribution during canvassing. Of the Northern Ireland workshop participants, Sarah Ewart is a litigant who has undertaken legal action to challenge Northern Ireland's abortion law.[6] Her mother Jane Christie, who also contributed to the conversation, accompanied her. Emma Campbell and Kellie O'Dowd were then co-directors of Alliance for Choice. Goretti Horgan is an activist with Alliance for Choice in Derry, work which she undertakes alongside her role as an academic at Ulster University working on abortion research.

Our aim in bringing these eight interlocutors together was not to 'tell the whole story' of abortion law reform 'North and South'; rather it was to begin a conversation on feminist activists as key actors in processes of legal change, complementing existing literature on 'formal' law reform in both jurisdictions.[7] We acknowledge that if eight different women had been in the room with us, or if the same eight women had been in the room with us six months earlier or later, the event could have been very different. We are conscious too that our participants did not include people working specifically to agitate for reproductive autonomy and freedom for women of colour, migrant women, women with disabilities, Traveller women, transpeople, and other specifically impacted persons who may seek abortion care. Those activists' experiences of law reform will have been different, and their experience of, engagement with, and perhaps rejection of feminist law work and its inevitable pragmatism is worthy of separate exploration. For the

6 *Ewart's (Sarah Jane) Application* [2019] NIQB 88.

7 For example, Joanna Erdman, 'Procedural abortion rights: Ireland and the European Court of Human Rights' (2014) 22(44) Reproductive Health Matters 22; Máiréad Enright et al, 'Abortion law reform in Ireland: a model for change' (2015) 5(1) feminists@law <https://journals.kent.ac.uk/index.php/feministsatlaw/article/view/173>; Fiona de Londras, 'Fatal foetal abnormality, Irish Constitutional law, and *Mellet v Ireland*' (2016) 24(4) Medical Law Review 591; Fiona Bloomer and Eileen Fegan, 'Critiquing recent abortion law and policy in Northern Ireland' (2013) 34(1) Critical Social Policy 109; Fiona Bloomer and Kellie O'Dowd, 'Restricted access to abortion in the Republic of Ireland and Northern Ireland: exploring abortion tourism and barriers to legal reform' (2014) 16(4) Culture, Health and Sexuality 366.

purposes of this starting conversation, however, these were the eight activists with whom we engaged, and whose experiences animate our account.

In Northern Ireland, Alliance for Choice is the leading grassroots voice on abortion law reform. In addition to providing direct practical support to women who need abortions, it engages with other advocacy groups, such as Amnesty International, the Northern Ireland Human Rights Commission, and the London Irish Abortion Rights Campaign. In Ireland, particularly in the lead-up to the referendum vote, a much larger collective action campaign came together which included a wide range of loosely connected groups and individuals; some single issue, and others more established. Groups such as the Coalition to Repeal the 8th Amendment were key to persuading organisations which do not have a primary focus on reproductive rights issues, such as trade unions, to enter the fray on behalf of the movement. During the formal referendum campaign, three organisations – the Abortion Rights Campaign, the Coalition to Repeal the 8th Amendment and the National Women's Council of Ireland – came together to form an official civil society campaign for a 'Yes' vote, called Together for Yes.[8] Most, although not all, pro-choice organisations in Ireland were affiliated with Together for Yes during the referendum campaign. In spite of these differences, the movements in both jurisdictions are interlinked and, in some instances, interdependent, and they have often worked together. Women from the North, and particularly from Alliance for Choice, campaigned actively for the 'Yes' vote in the Republic's referendum, especially in the border counties. Similarly, since the Amendment was repealed, some activists in the South have been involved in work to facilitate access to abortion services for women from Northern Ireland. Activists are conscious that women who live in one jurisdiction may, in future, need to access services in the other. For these reasons, it is productive to consider all of these activists together.

By use of the term 'activist', we mean social change agents who participate in some form in a social movement/collective action.[9] Self-identifying as an activist is not essential to this definition,[10] and indeed some of our participants did not overtly embrace this label. Our engagement with feminist activists in the latter stages of reform struggles on both parts of the island makes clear the importance to law and lawyers of recognising how these activists' actions in, with, to and against law helped to shape reform processes. In engaging with our activist participants, we were not aiming to arrive at 'the truth' of feminist activist engagement in law in the abortion rights movements in Ireland and Northern Ireland; any attempt to do so would be as distorting[11] as attempts to create dominant narratives of abortion law reform that underplay or even invisibilise feminist, activist labour. Neither do we attempt to narrate the internal dynamics of these groups

8 For official accounts of Together for Yes's strategy and activities, see Michael Barron, *Learning from the 2018 Together for Yes Campaign* (2019) <https://www.togetherforyes.ie/app/uploads/2019/11/2019_TFY_Review.pdf?fbclid=IwAR33cJUSgdH-AUFCqbh3Qh01oCLk8Sk0s2EqHanZj2e7VejHiV0SfEJ_k3U>

9 Chris Bobel, '"I'm not an activist, though I've done a lot of it': Doing activism, being activist and the 'perfect standard' in a contemporary movement' (2007) 6(2) Social Movement Studies 148. For wider discussion, see Leila Rupp and Verta Taylor, 'Forging feminist identity in an international movement: a collective identity approach to social movements' (1999) 24(2) Signs 363; Jennifer Baumgardner and Amy Richards, *Manifesta: Young Women, Feminism, and the Future* (Farrar, Strauss & Giroux 2000); Scott Hunt and Robert Benford, 'Collective identity, solidarity and commitment' in David Snow, Sarah Anne Soule and Hanspeter Kriesi (eds), *The Blackwell Companion to Social Movements* (Blackwell 2004).

10 Bobel ibid 157.

11 See generally Antoinette Burden, '"History" is now: feminist theory and the production of historical feminisms' (1992) 1(1) Women's History Review 25.

or the many other feminist groups engaged both recently and earlier in the reproductive rights struggles in Ireland. We recognise that there is no singular truth of these engagements, or one story of reform; rather there are multiple stories and narratives that together form the complex, textured background to the reform of the law in both jurisdictions. However, there is a 'past' to which multiple actors increasingly seek to lay claim; notably, in the South, some political parties seek to represent repeal of the 8th Amendment as a product of their actions or of a 'quiet revolution',[12] seemingly erasing the substantial feminist law work that underpinned it. Informed by the intellectual commitments of feminist historiography,[13] we know that recognising the authority of feminist activists' narrative accounts is part of creating the legal history of abortion law reform in Ireland.[14] As Cheryl Glenn has written in another context, 'in choosing what to show, how to represent it, and whom to spotlight … [we] subtly shape our perceptions'[15] of the past. In this paper, we choose to show, represent, and spotlight feminist law work and the activism that underpins it. Concentrating on these narratives makes clear the importance of rejecting any simplistic 'legal' telling of abortion law reform in Ireland and Northern Ireland in which court cases, legislative reform, or parliamentary debates are presented as isolated and self-generated incidences of legal and political engagement.

Although it is not an oral history or biographical project, our ethical and methodological approach to the material offered to us by the participants is informed by a commitment to pluralising the historiographical record(s) of significant events, and to the 'ethically and intellectual responsible gesture that disrupts … frozen memories in order to address silences, challenge absences, and assert women's contributions to public life'.[16] As such we do not fact-check or authenticate the accounts and recollections of our participants; nor is any such process necessary. Instead, we recognise personal recollection and experience as a valid source of understanding processes of legal change and see the collection of such accounts and their analysis as a means of understanding narrated memories to which law and lawyers are generally insufficiently attentive. In this, we are informed and influenced by feminist commitments to recognising the authenticity of first-person narrative, the epistemological value of personal experience, and the standing of activists as the most qualified narrators of their own participation in activist endeavours.[17] Thus, while we contextualise the workshop and the moment in time in which we spoke to these feminist law workers in Part 2, we do not otherwise seek to explain, validate, or somehow 'authenticate' participants' accounts. Instead, we recognise them as the experts in their own experiences of feminist law work.

We use the word 'work' for a reason. This paper talks about a range of legal activities which formed part of everyday abortion rights activism; peer and public legal education,

---

12 Following the result of the referendum to repeal the 8th Amendment, Taoiseach Leo Varadkar claimed that 'a quiet revolution has taken place'. See, for example, RTE News, 'Result is culmination of quiet revolution, says Varadkar' 26 May 2018 <https://www.rte.ie/news/eighth-amendment/2018/0526/966132-reaction/>.

13 Following Chandler, we understand historiography as 'a critical consciousness at work in the writing of history': James Chandler, *England in 1819: The Politics of Literary Culture and the Case of Romantic Historicism* (University of Chicago Press 1998) 77.

14 As Enoch and Bessette write, 'Feminist recovery … depends on the archive.' Jessica Enoch and Jena Bessette, 'Meaningful engagements: feminist historiography and the digital humanities' (2013) 64(4) College Composition and Communication 634, 637.

15 Cheryl Glenn, 'Lies, and method: revisiting feminist historiography' (2000) 62(3) College English 387, 388.

16 Ibid 389.

17 See generally the contributions to Sandra D Harding (ed), *The Feminist Standpoint Theory Reader: Intellectual and Political Controversies* (Routledge 2004).

developing new legal skills, lobbying institutional legal actors, mobilising popular support for concrete legal change, generating narratives to delegitimate existing law and justify reform, organised law-breaking, supporting litigation and more. These projects were multifaceted: activists were doing many things at once, often across many different institutional, political and community environments, and recognising that different dispositions towards law are required and appropriate in different settings. They might be drafting legislative amendments in one environment and distributing 'illegal' (but safe) abortion medication in another. These activities were often unpaid, considered unprofessional or illegitimate, done 'on the move', or in private or domestic space. In characterising these activities as 'work', we draw on a long feminist tradition of bending and stretching the concept of work in order to honour women's marginalised labour.[18] We not only recognise activists' collective and individual engagement with law as work, but demonstrate that aspects of mainstream law reform work are dependent on feminist law work At the same time, we recognise that feminist law work is alienated labour[19] – it is not only a site of value and meaning-making. It can also be difficult and disappointing. As we see in this paper, feminist law work does not indicate a legal vocation, or an unambivalent commitment to law. It often looks more like reluctant, resistant or pragmatic engagement.

Our attention to feminist law work on the island of Ireland has relevance beyond these jurisdictions. What this micro-story reveals is that feminist engagement with law, even if 'cheeky'[20] and irreverent, can be productive and heterodox, challenging conventional or dominant tropes about legal (im)possibility, and that activists engaging as feminist law workers consciously reclaim law from 'elite' disciplinary or professional 'ownership' by developing skills and knowledge that allow them to trouble, and force change within, orthodox legal thinking. Thus, while presenting a 'snapshot' of the abortion law reform processes in Ireland and Northern Ireland, this paper also identifies and insists upon the importance of a frame for engagement – this feminist law work – that is discernible in reproductive rights activism all over the world, as well as in other social justice campaigns in which feminist activism is at the forefront.

2 Contextualising abortion activism in Northern Ireland and Ireland

As already noted, abortion laws in the Republic of Ireland and Northern Ireland have recently seen legal transformations unimaginable even a decade ago. In Northern Ireland, abortion was previously criminalised under sections 58 and 59 of the Offences Against the Person Act 1861 with very few exceptions. In October 2019, abortion was decriminalised in the jurisdiction by the Northern Ireland (Executive Formation etc) Act 2019. This was legislative action taken by Parliament at Westminster during a period where the devolved Assembly at Stormont was not functioning. Regulations pertaining to legal provision of abortion were introduced in March 2020. In the Republic, the

18 See further, Kathi Weeks, *The Problem with Work* (Duke University Press 2011); Silvia Federici, *Revolution at Point Zero* (PM Press 2012); J K Gibson-Graham, *A Postcapitalist Politics* (University of Minnesota Press 2006).

19 Alison Jaggar, *Feminist Politics and Human Nature* (Rowman & Littlefield 1983) 307

20 In this we are influenced by Fletcher's articulation of 'cheeky witnessing' which she claims has three elements: 'First, it is messy and irreverent in innovating with names to display the mixed genealogies of sources of feminist knowledge. Second, cheeky witnessing generates novel subject-figures, such as migrant cleaners, who make knowing connections between different reproductive labourers as observers of the trail in diaspora space. Third, cheeky witnessing places funny objects, knickers in this instance, so as to join up particular public locations and make them more, if unevenly, comfortable for sexual and reproductive bodies.' Ruth Fletcher, 'Cheeky witnessing' (2020) 124 Feminist Review 124.

constitutional provision which prohibited almost all abortions – Article 40.3.3, known as the 8th Amendment – was repealed in May 2018. New abortion legislation, the Health (Regulation of Termination of Pregnancy) Act 2018, was passed some months later and, in principle, abortion care is now legally available across the country.

This change was a long time coming. Neither jurisdiction embraced liberalising action to reform the law on abortion for several years.[21] Institutional dynamics within each jurisdiction mean that change to the law took different forms. In Northern Ireland, for example, a referendum was unnecessary since legal change was largely in the hands of elected representatives. Targeted strategic litigation has recently been employed in domestic courts to some positive effect.[22] The relative institutional autonomy of Westminster and the devolved Assembly at Stormont facilitates change to abortion law by the ordinary mechanisms of legislative reform, should elected representatives be moved to use them.[23] This situation was, however, complicated by the collapse of the Stormont Assembly between January 2017 and January 2020, and it was eventually Westminster legislation that facilitated decriminalisation. In the Republic, by contrast, the 8th Amendment and an unreceptive judiciary were perceived to present an impenetrable obstacle to successful litigation. Strategic cases and petitions were brought in international spaces instead. A referendum was not only unavoidable but perceived as a beneficial way to break legislative deadlock.

There are thus important differences in experiences across the jurisdictions. However, there are also inevitable commonalities and connections. First, both jurisdictions have had definitively non-linear experiences of the road to reform of abortion law. In these non-linear processes, no one event in either jurisdiction was responsible for triggering reform. Rather, legal change followed the culmination of multiple events particular to each jurisdiction. Contingent, and often unexpected, moments have been opportunities to develop and publicise reform work. In the Republic, these moments include the death of Savita Halappanavar in 2012 and the experiences of Amanda Mellet and Siobhan Whelan in 2010 and 2011 who were required to travel to access abortion services in cases of fatal foetal abnormality. In Northern Ireland, Sarah Ewart's sharing of her experience of fatal foetal abnormality in 2013, a judicial review by the Northern Ireland Human Rights Commission initiated in 2015,[24] and an Inquiry by the United Nations Committee on the Elimination of All Forms of Discrimination Against Women (CEDAW) which reported in 2018[25] are all notable examples. Second, legal change is only the starting point for wider processes of transition in Ireland and Northern Ireland. The new abortion laws are still being embedded and reconfigured in both parts of the island. Details around practical provision and application of the law remain to be settled. In this respect, reform was not the end of the story. At the time of our interviews and workshop activists were still working on influencing the practical details and roll-out of the new legislation on the

21 Siobhan Mullally, 'Debating reproductive rights in Ireland' (2005) 27(1) Human Rights Quarterly 78; Goretti Horgan, 'Abortion and citizenship rights in a devolved region of the UK' (2014) 13(1) Social Policy and Society 39; Lisa Smyth, *Abortion and Nation: The Politics of Reproduction in Contemporary Ireland* (Routledge 2017); Sheldon et al (n 4).

22 *Family Planning Association for Northern Ireland v Minister for Health, Social Services and Public Safety (SPUC NI and others intervening)* [2004] NICA 39; *The Northern Ireland Human Rights Commission's Application* [2015] NIQB 102.

23 Jennifer Thomson, *Abortion Law and Political Institutions: Explaining Policy Resistance* (Palgrave 2019).

24 *Northern Ireland Human Rights Commission* (n 22).

25 CEDAW Committee, Report of the Inquiry Concerning the United Kingdom of Great Britain and Northern Ireland under Article 8 of the Optional Protocol to the Convention on the Elimination of All Forms of Discrimination Against Women (2018) UN Doc CEDAW/C/OP8/GBR/1.

ground. Third, feminist activists North and South of the border were motivated to speak back to the effects of law on the lived experiences of gendered subjects and troubling restrictive and heteronormative legal frameworks.[26] They were concerned with the empowerment of women and pregnant persons, attentive to long-term, structural gender imbalances within their respective jurisdictions and affected by first-hand experience of the effects of the law on the ground. We will now turn to explore the work they undertake in more detail.

3 Ambivalence about law in feminist law work

Despite law's obvious uses – its symbolic weight[27] and its connection to coercive power that can be used to redistribute rights and resources – feminist organising has always had an ambivalent relationship with law; feminists cannot help but be conscious of law's violence.[28] There are recognised dangers; that a movement will invest too many of its resources in legal change and, in the process, become complicit with darker aspects of legal power.[29] Law tends to subject women to ever more extensive legal control, even in the guise of liberalisation, and can impose its logic on wider political discourse in ways that suppress women's desires and experiences.[30] In discussions with our participants, it became clear that, while they engaged closely with law and its various spaces, they maintained a similar feminist suspicion of law. This was evident in how they understood the function of law in their activist campaigns for change. The activists in our conversation were clear that they undertook efforts towards reform because they knew, as Emma Campbell noted when we spoke with her, that the prevailing law harmed women and saw advocating for legal change as an essential 'harm reduction exercise'. They were not, however, comfortable with law. In part, this was because law's insiders are not always comfortable with them, or with their place as agitators for law reform. Goretti Horgan brings up this issue in the context of the relationship between feminist protest and litigation:

> So, we always had this kind of thing where we were talking about legality even as we were street campaigning, you know? We never saw the two as being separate. I think that is until the *Family Planning Association case*.[31] And then we [activists] were told that we shouldn't do any campaigning until the case was out of the way, until the case was heard and ruled on. Because you know, were we to be on the streets or we were told we shouldn't be outside of the court, you know, we shouldn't be doing any campaigning really, you know just let the courts handle it. Now, obviously we didn't take any notice of that. But there was, like an imperative for a while there where you know, where we were frowned upon, where we were seen as being kind of wreckers.

26 For an overview of how feminist legal scholars have described and approached these types of frameworks, see Katharine Bartlett (ed), *Feminist Legal Theory: Readings in Law and Gender* 2nd edn (Routledge 2018); Janet Halley, Prabha Kotiswaran, Rachel Rebouché and Hila Shamir (eds), *Governance Feminism: Notes from the Field* (University of Minnesota Press 2019).

27 Sally Engle Merry, 'What is legal culture? An anthropological perspective' (2010) 48(2) Journal of Comparative Law 40.

28 Austin Sarat and Thomas Kearns (eds), *Law's Violence* (University of Michigan Press 1993); Rosemary Hunter, 'Law's (masculine) violence: reshaping jurisprudence' (2006) 17 Law and Critique 27; Hilary Charlesworth, 'Feminist ambivalence about international law' (2005) 11 International Legal Theory 1.

29 See, for example, Ratna Kapur, 'Human rights in the 21st century: take a walk on the dark side' (2006) 28(4) Sydney Law Review 665.

30 Nivedita Menon, *Recovering Subversion: Feminist Politics beyond Law* (University of Illinois Press 2004) 5.

31 *Family Planning Association* (n 22). For discussion see Ruth Fletcher, 'Abortion needs or abortion rights? Claiming state accountability for women's reproductive welfare' (2004) 13(1) Feminist Legal Studies 123.

In their ambivalence towards law, feminist activists refuse law's respectability, decentre law in their work and campaigns, and often engage in a creative and provocative play around law. Let us explore each of these in turn.

3.1 Refusing law's respectability

At a fundamental level, feminist activists refuse the respectability of law and the formal legal domain. In other words, they question its authority and proprietary place in social life. Of course, that respectability is complicated. In the Northern Irish context, more broadly, activists are conscious of an older history of law's role in legitimating or not effectively counter-acting violence.[32] Emma Campbell summarised:

> And, you know, growing up in Northern Ireland, the law was always very much used as a tool of control to the people as far I understood it, you know. We had, you know, we grew up having to go through army turnstiles to get into the town centre every day, that's how we, how I feel that I saw the expression of the law in a, I don't know, embodied in a human form. So, I don't give a shit about the law, really [laughs].

This lack of automatic respect for law has played out in more recent pro-choice organising in both jurisdictions. While law's failings encouraged feminist activists to engage in feminist law work towards reform, these same failings meant that law was marginalised when activists imagined their desired outcomes. For example, feminist organisations such as Alliance for Choice and the Coalition to Repeal the 8th Amendment did not, in principle, favour the creation of new legislation to regulate abortion access. They preferred the approach of jurisdictions such as Canada and Cuba where regulation is the role of the healthcare, as opposed to the legal or criminal justice, system.[33] For these groups, the preference was repeal of the existing law, followed by decriminalisation and limited legal regulation. As Ailbhe Smyth put it, 'none of the pro-choice organisations really wanted there to be legislation'. However, this preference was gradually tempered by an understanding that law – and feminist engagement with it – was necessary in the given circumstances.

In the Irish context, there seemed to be an understanding that moving from hyper-restrictive constitutional regulation to 'no' statutory regulation of abortion would not garner sufficient support, at least in political institutions, so that during the referendum pro-choice organisations endorsed the government's proposed legislative scheme, albeit at times with some reservations. For Linda Kavanagh, the sense of law's symbolic and communal role was also crucial:

> To be honest … the legality of abortion isn't necessarily that important to your grassroots anarchist leftie activist. But in the wider sense, we know that people love the law. They think that morality lies within the law or some of their morality comes from the law, which to me is really bizarre but is true. So that when you decriminalise something, you do remove some of the stigma and shame. And so, kind of being aware of that. So, that's kind of where respectability comes in there.

32 See, for example, Fionnuala Ní Aoláin, *The Politics of Force: Conflict Management and State Violence in Northern Ireland* (Blackstaff Press 2000); Fionnuala Ní Aoláin and Michael Hamilton, 'Gender and the rule of law in transitional societies' (2009) 18(1) Minnesota Journal of International Law 380; Monica McWilliams, 'Violence against women and political conflict: the Northern Ireland experience' (1997) 8 Critical Criminology 78.

33 Rachael Johnstone, *After Morgentaler: The Politics of Abortion in Canada* (University of British Columbia Press 2017); Carrie Hamilton, *Sexual Revolutions in Cuba: Passion, Politics, and Memory* (University of North Carolina Press 2017).

This account of the emergence of an unsettled, if strategically coherent, attitude to law and its respectability is not surprising. Patricia Ewick and Susan Silbey argue that legal consciousness is 'multiple and contingent'; simultaneously obedient and resistant.[34] In the same situation, individuals may maintain a sense of themselves as 'before', 'with' and 'against' the law. It is not surprising then that participants in our workshop maintained a strong critical oppositional sense of law's dualities and were reluctant to embrace legal change as an end in itself, even as they engaged law sometimes for their own purposes.[35] In feminist law work, this refusal of law as something respectable and deserving of deference is accompanied by a decentring of law.

3.2 Decentring law

Carol Smart famously encouraged feminists to decentre law and prefer other strategies for change in order to avoid legal colonisation of feminist campaigning.[36] This was echoed in our participants' clear sense that legal change was not sufficient to deliver meaningful reproductive justice. Irish and Northern Irish feminists are under no illusion that law offers an easy solution to the problems they are grappling with. A reluctance to engage with law as an end in itself appears to stem from participants' concrete experience of legal frameworks that have either withheld help from or failed appropriately to respond to the real-life needs of women and pregnant persons. This reluctance does not dissipate even when some of the key demands and desired outcomes of feminist law work have been achieved, such as repeal of the 8th Amendment. The long delay between the referendum date (25 May 2018) and the coming into force of the legislation that made provision for lawful abortion (1 January 2019) – the time during which our workshop took place – was a temporal site of real frustration with the law. Erin D'Arcy reflected on this, saying:

> I just received [an online message] last night … it's been six months since we had the vote and … she's having to email [the 'In Her Shoes' Facebook page] anonymously instead of just going to her GP.

In some ways, these disappointments are anticipated from the beginning of feminist law reform campaigns. As Kellie O'Dowd said of Alliance for Choice, 'we realised very early on in our campaign that legislative change is never going to be enough'. This insight was not only a product of experience in abortion rights activism but a conclusion drawn from feminist law work in other areas.[37] For her, campaigns to achieve equal pay in the 1970s in particular revealed that formal law reform is often limited in what it can achieve and must be supplemented with meaningful change in socio-political attitudes and culture. Legal struggle is intertwined with social struggle, and feminist demands extend beyond anything law alone could provide.[38] This understanding leads legislative reform to be one, but not the only, focus of the work Alliance for Choice does. As O'Dowd put it:

---

34 Patricia Ewick and Susan Silbey, *The Common Place of Law: Stories from Everyday Life* (University of Chicago Press 1998).

35 Michael McCann, *Rights at Work: Pay Equality and the Politics of Legal Mobilization* (University of Chicago Press 1994) 233.

36 Carol Smart, *Feminism and the Power of Law* (Routledge 1989).

37 For discussion, see Reg Graycar and Jenny Morgan, 'Law reform: what's in it for women?' (2005) 23(2) Windsor Yearbook of Access to Justice 393; Susan Armstrong, 'Is feminist law reform flawed? Abstentionists and sceptics' (2004) 20(1) Australian Feminist Law Journal 43.

38 Boaventura De Sousa Santos, 'The World Social Forum and the global left' (2008) 36(2) Politics and Society 247.

> … along with campaigning and lobbying around legislative change, we need to change the hearts and minds of people … we bring a very grassroots and personal experience to law reform but we don't put all our eggs in one basket.

Similarly, Linda Kavanagh remembers that the stakes the Irish referendum were far broader than constitutional change:

> And I suppose, no, I don't feel like I was just advocating for constitutional reform, and very few people that I would ever deal with would be in it simply for the constitutional reform aspect. It was that we knew we needed to make this change. We knew it was … For me and for the organisation that I volunteer with, it's that we knew this was the next obstacle and we weren't going to get anything unless this was removed … But I think not a lot of people would just be in it for the sheer joy of change of the constitution.

Awareness of this relationship between law and social change does not necessarily lead to the same strategic conclusions. For Kellie O'Dowd, awareness of the limitations of the incrementalist approach to legislation then about to be passed in the Republic only strengthened the need to ensure that the Alliance continued to ask for the broadest and most comprehensive possible legislative provision in the North. By contrast, for Ailbhe Smyth, in heading a coalition whose members differed widely on the proper legal approach to abortion, and in the absence of a legislative proposal that all were agreed on, acknowledging law's limitations meant adopting a narrow legal project. She explained that legal change takes place in stages, and her role, as she saw it, was necessarily confined to getting the initial vote in favour of repeal of the 8th Amendment, before returning the legislative question to the parliament.

> I saw my job in the Coalition as repealing the Eighth Amendment – actually, we did that, and that had to be done … I don't think any of us ever thought … or most of us didn't think that the legislation we would get straight up after repeal would be perfect, and indeed it is obviously very, very far from perfect … I'm not the expert on this aspect of things. So I'll quite happily say, not that it's altogether in other people's hands, because we all have a vested interest but that the legislative part is really in your [ie lawyers'] hands now, more than in the hands of activists like me.

Although they were aware of law's limitations, this group of activists was also conscious of its power. Although law was not venerated, it could be used instrumentally to try to compel action by otherwise-unresponsive states. They could deploy law as 'a stick to beat the establishment with' (Linda Kavanagh). They recognised the formal, establishment power of the law to compel the state. At the same time, this recognition was tinged with irreverence for law and for state power. As Linda Kavanagh put it, engagement with law is such an effective tool of persuasion for the establishment because, 'you know … they are horny for the law. They love it.' For Alliance for Choice, refusal to venerate law was bound up with a willingness to get around it, ridicule it, break it, or subvert it as necessary to ensure abortion access for the women they wanted to support. Emma Campbell noted:

> … as an activist organisation we've been happy not to take the law that seriously on many occasions … the law is an ass so we'll do whatever we can around, above and beyond it, whenever we can.

3.3 Tickling law

Precisely because it does not take its law 'that seriously', Alliance for Choice has been able, thirdly, to emphasise creativity in their approach to law reform. It is less persuaded by law's existing limits and boundaries than by the need to respond to problems on the

ground. This has led feminist activists to be more interested in innovatively combining international approaches than establishment lawyers might be. Emma Campbell explains:

> [I am a] creative problem solver because I come from, you know, a creative background. And I get a little bit frustrated when legal people, let's say, or political people use or describe or imagine what the future looks like but only looking at their small tiny bit of experience or example of what they already have in front of them. Because, you know, especially from being at international conferences and you see the brilliant way that some countries have dealt with the fact that they have restrictive laws but they have interpreted, for instance, they have interpreted the word 'health' to include the World Health Organisation interpretation of that which is, which includes social issues, which includes mental health, which includes being able to provide for your family. So ... on a day to day basis we get people phoning asking us how we can help them. And so as people who have been at the, I guess the sharp end of physically providing abortion help or providing abortions or physically being the body in between a protestor and a person, then your approach is probably a little bit more pragmatic and a little bit less … [W]e feel frustrated that they so almost pedantically interpret the law, interpret the law conservatively and don't see any wiggle room in there where others might ... you know, there are definitely lawyers in other countries and in other contexts that have the same law in front of them and see it differently, you know.

Emma describes Alliance for Choice's feminist law work as 'tickling law'; that is, deploying agentic resources to go beyond a docile and disciplinary relation to law in order to engage in play and provocation around law and its limits.[39] In this respect, feminist law work often employs significant creativity, playing games with law and pushing its boundaries in ways that transcend law's control.[40] This is designed to test law, to reveal its hidden limits, effects or hypocrisies and, ultimately, to stimulate change. Emma reflects on how tickling has been carried out through tactics like covert distribution of abortion medication, women publicly declaring that they have procured or helped to procure an abortion, and openly providing information on the ways in which feminist activists assist in pushing the boundaries of the criminal law.[41] This strategy of 'tickling' serves to stage a disruption to law, legal time and its subjects. It challenges law's dominant narrative about itself in acts of resistance. These acts are reminiscent of the actions of others who have acted as provocateurs to law in areas such as welfare,[42] racial injustice[43] and decolonisation.[44]

39 On criminality as provocation in struggles for law reform in Ireland, see Máiréad Enright and Emilie Cloatre, 'Transformative illegality: how condoms "became legal" in Ireland, 1990–1993' (2018) 26(3) Feminist Legal Studies 161; Emilie Cloatre and Máiréad Enright, '"On the perimeter of the lawful": enduring illegality in the Irish family planning movement, 1972-1985' (2017) 44(4) Journal of Law and Society 471.

40 A similar approach is evident among some artistic interventions in abortion law reform in Ireland: Fletcher (n 31).

41 Radhika Sanghani, '"Arrest us": Northern Ireland women want to be prosecuted for breaking abortion laws' *The Telegraph* (London 26 June 2015) <https://www.telegraph.co.uk/women/womens-life/11700651/Abortion-Northern-Irish-women-want-arrest-over-illegal-abortion-pills.html>.

42 Austin Sarat, '"…The law is all over": power, resistance and the legal consciousness of the welfare poor' (1990) 1(2) Yale Journal of Law and the Humanities 343.

43 Patricia Ewick and Susan Silbey, 'Conformity, contestation, and resistance: an account of legal consciousness' (1991) 26(1) New England Law Review 731.

44 Mindie Lazarus-Black and Susan Hirsch (eds), *Contested States: Law, Hegemony and Resistance* (Routledge 1994).

4 The multifaceted nature of feminist law work

The participants in our workshop and interviews are neither lawyers nor institutional feminists employed by state or government bodies.[45] However, their perspectives on legal change should not be viewed as 'external'; they have been intimately caught up in legal change. In particular, they are adept at identifying locations of law work and at strategising across various locations – formal and informal, parliamentary and judicial, national and transnational – to build momentum towards formal legal change. In consequence, the locations of feminist law work are multifaceted and overlapping.

Given the long-term lack of movement towards reform via referendum (Ireland) or legislative change (Northern Ireland), feminist activists in both jurisdictions had to find alternative fora and spaces – domestic courts and international human rights bodies, for example – in which to develop and publicise their arguments for change. In other words, when legislatures were unresponsive to demands for a liberalising referendum,[46] more permissive constitutional interpretations,[47] or legislative change,[48] feminist activists mobilised other legal sites for 'politics by other means'.[49] Legal arguments developed at these sites were then translated back into potential legislative discourse once the opportunity emerged, sometimes with a view to compelling parliamentarians to act, or reacting to politicians reneging on commitments to support change. Speaking of her daughter Sarah Ewart's experiences in trying to secure help from politicians after having to travel for termination in a case of fatal foetal anomaly, Jane Christie captured this process and strategy of moving between politics and law:

> … we wrote to 108 MLAs ... we ended up trying our best as Sarah said for two years going round the MLAs, we got absolutely … there was plenty of understanding, plenty of comparisons in their own situations but nobody was actually listening to us. We got agreement from the DUP that they would not use a petition of concern and they would support [a Bill proposing access to abortion in cases of fatal foetal abnormality]. At that stage then they back-pedalled with us, they had agreed also a change in law, [a senior politician] back-

45 Holly McCammon and Amanda Brockman, 'Feminist institutional activists: venue shifting, strategic adaptation, and winning the Pregnancy Discrimination Act' (2019) 34(1) Sociological Forum 5.

46 In Ireland the referendum was not confirmed until after both a Citizens' Assembly and a specially convened Oireachtas Committee recommended it; Cabinet did not propose it to the Oireachtas spontaneously. The only substantive referenda on abortion that had been proposed since 1983 had, in fact, sought further to restrict access to abortion and to reverse the Supreme Court decision in *Attorney General v X* [1992] 1 IR 1 that abortion was constitutionally permissible where the pregnant person was at real and substantial risk of suicide. See further Máiréad Enright, 'Abortion law in Ireland: reflecting on reform' in Lynsey Black and Peter Dunne (eds), *Law and Gender in Modern Ireland: Critique and Reform* (Hart Publishing 2019) 55.

47 In Ireland attempts to pass legislation that would permit abortion in cases where the foetus had been diagnosed with a fatal anomaly and which were arguably permissible given a more liberal constitutional interpretation were unsuccessful, based largely on Attorney General advice that any such legislation would be unconstitutional. This is in spite of the fact that this was far from entirely settled as a matter of constitutional interpretation. See further Jennifer Schweppe and Eimear Spain, 'When is a foetus not an unborn? Fatal foetal abnormalities and Article 40.3.3' (2013) 1 Irish Journal of Legal Studies 92; Aisling McMahon and Bríd Ní Ghráinne, 'Reconsidering the Irish Constitution and international law: abortion and fatal fetal abnormality' (5 March 2018) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3134542>.

48 When Ireland finally legislated to give effect to the limited right to access abortion under the 8th Amendment in 2013, the legislation was unduly restrictive and all attempts to amend the proposed legislation so that it was compliant with the Constitution but not punitive failed. See Claire Murray, 'The Protection of Life During Pregnancy Act 2013: suicide, dignity and the Irish discourse on abortion' (2015) 25(6) Social and Legal Studies 667.

49 Richard Abel, *Politics by Other Means: Law and the Struggle against Apartheid (after the Law)* (Routledge 2002).

> pedalled with us, and at that point Sarah and I realised that to get any help we were going to have to go through the courts.

This comment also gestures towards the first location of multifaceted feminist law work: domestic courts and litigation.

4.1 DOMESTIC COURTS

In Northern Ireland domestic litigation was a core element of the strategy for legal mobilisation and law reform. This took a number of forms and particular recent high-profile cases – Ewart's case,[50] *A and B*[51] and the judicial review initiated by the Northern Ireland Human Rights Commission[52] – were turning points, not only in terms of substantive law but also in terms of connecting legal doctrinal abstraction to empirical lived realities for the electorate. Speaking directly to Sarah Ewart, Kellie O'Dowd captured the importance of her case for the movement quite beyond the strict legal outcome:

> … until Sarah, until what happened to you did happen and came out into the public, like, you know, we would have meetings and there'd be two people and a dog turn up, and you're kind of going abortion law reform's never going to happen, we're never going to get the [19]67 Abortion Act extended. But then when your story went public, it absolutely changed the discourse because if you have a situation where it's surrounded by silence and stigma, but you were the human face and you were the wanted, the very much wanted pregnancy and it kind of really opened people's minds to going 'what do you mean she can't access abortion here? What do you mean she has to be basically, she has to go to London to access this?' That made people interested and it made them sit up so I know that youse [referring to Sarah and Jane] didn't know what was happening in terms of what was going on for you, but as an abortion rights activist who had been working on this for five years, it was the light bulb that we had needed in terms of saying this is what happens to women who can't access reproductive healthcare here. So, I want to thank you … And I always said that … the tipping point when we're out doing workshops would be the Sarah Ewart case and everybody knows that case and everybody can relate to that and it's so tangible because it allows us to have those conversations.

This seemed in some ways to be precisely what Sarah Ewart hoped would be the effect of her decision to pursue justice through the courts. Even though her mother, Jane Christie, said '[w]e don't class ourselves as activists but we can clearly see that there is a problem here in Northern Ireland … and it's disgusting', Sarah's motivation was clear. In her words, she decided 'we'll just keep going until we try and get it changed'. This illustrates well the agency and the feminist law work involved for individual litigants in taking cases of this kind: Sarah's case was taken partly for herself, but partly also to redress an injustice that was systemic and that was impacting on other women, and which politicians in Northern Ireland did not seem to be willing to resolve.

Other cases pursued not by those seeking reform, but by the state, also played important roles in turning the tide in Northern Ireland; particularly the case of a woman

50 *Ewart* (n 6).

51 R *(on the application of A and B) (Appellants) v Secretary of State for Health (Respondent)* [2017] UKSC 41.

52 *Northern Ireland Human Rights Commission* (n 22); *In the Matter of an Application by the Northern Ireland Human Rights Commission for Judicial Review (Northern Ireland)* [2018] UKSC 27.

who was prosecuted for acquiring abortion medication for her daughter.[53] It reinforced – North and South – the potency of continuing criminalisation, and activists were involved in ensuring that, as much as possible, the case proceeded through the courts rather than disappearing quickly through a guilty plea. Goretti Horgan reflected on this case to show how this process illuminated both the inadequacy of the existing law's framing of abortion, and the lived costs of enforcing the law:

> That actually makes me think … [of] the mother who's taken the judicial review on the decision to prosecute her for getting the abortion pills for her daughter, because when that was reported in the media, it was reported that she was charged with getting 'poison' to try to cause her daughter to have an abortion.[54] Now the poison was those very safe pills and that are available on the NHS, et cetera. And … everybody else pleaded guilty and accepted a caution or whatever so there was never any, there was never any narrative about them except what we ourselves kind of knew in private or whatever … And I think that some of what's come out since, I mean the fact that the police actually took the daughter out of her classroom, to be interviewed … we wouldn't have heard that if the court case hadn't happened. You know, there's a whole lot of like small details like that that have absolutely horrified even people who are against abortion. They're saying why would they do that to a wee girl who's had enough on her plate anyway. People who were forced, there have been too many people who were forced to tell their stories … You know, people forced to tell their stories who maybe really would want – including you, Sarah, you know, who'd prefer to be private and everything but are forced to tell them because of these awful laws that we have and to go through court cases that really, they shouldn't have to.

Feminist law work moves between locations – courts and legislature, for example – because it understands that legal principles formulated at one site may be transformed and transformative when they are translated for another. For instance, a case that has 'failed' in court may produce unexpected victories when it is considered by the legislature.[55] Goretti Horgan describes an example of the latter in the aftermath of the UK Supreme Court decision in *A and B*.[56] The litigants in that case asked for free access to abortion from the National Health Service for women resident in Northern Ireland. While the court ultimately decided against those women, Alliance for Choice's campaigning around this case was instrumental in the government's decision, announced the following day, to facilitate free access for Northern Irish women. That decision had a real and positive effect in making abortion accessible – albeit subject to travel – for many more women in Northern Ireland. Again, that was made possible through a long-term engagement with the question of securing free NHS abortions for people in Northern

53 Alan Erwin, 'Prosecution of Northern Ireland mum for buying daughter abortion pills "did not breach human rights," court rules' *Belfast Telegraph* (16 December 2019) <https://www.belfasttelegraph.co.uk/news/northern-ireland/prosecution-of-northern-ireland-mum-for-buying-daughter-abortion-pills-did-not-breach-human-rights-court-rules-38788749.html>.

54 This is a reference to section 58 of the Offences Against the Person Act 1861.

55 Michael McCann and George I Lovell, 'Towards a radical politics of rights' in Paul Christopher Grey (ed), *From the Streets to the State: Changing the World by Taking Power* (SUNY Press 2010) 139,155; Douglas Nejaime, 'Winning through losing' (2010-11) 26 Iowa Law Review 941.

56 *A and B* (n 51). For wider discussion, see Leanne Cochrane, 'Devolution and discrimination between citizens under Article 14 ECHR: preserving local provision' (2017) 76(3) Cambridge Law Journal 472; Stephen Cragg, 'Abortion, Northern Ireland and the NHS in England: can respect for devolved governments be a justification for discrimination?' (2017) 4(2) Journal of International and Comparative Law 377.

Ireland, and through productive engagement between activists and sympathetic practising lawyers. Goretti Horgan describes it thus:

> Whenever Alliance for Choice first started, and we were looking to get the Abortion Act [1967] extended to here … [we were told] it couldn't happen, it would destroy the Peace Process. Literally for ten years we were told that from 1998 to 2008 ... How, I don't know. So, then when we realised that wasn't going to happen, we started to look for free abortions on the NHS and I think like it was about 2002 we thought about taking a legal route and we started looking, we actually put out a press release and it was covered in the papers and all, looking for somebody who would be willing to try to get an NHS abortion and then take a case about it … and we were getting questions asked in Parliament all the time … we had lawyers in London … kind of older people, like my age, and they were keeping an eye the whole time and in 2012 they came and said, you know, you don't have to have somebody who's asked, we've checked, you don't have to have somebody who's asked for an NHS abortion, you just have to have somebody who's got like a good story and we'll take, you know, we will run. So, that's how the *A and B* case[57] came about. Now, we never thought that that would work and sure enough it didn't because actually they lost, they lost in the Supreme Court and yet the very next day, the announcement was made that there would be free NHS abortions … And certainly without the court case, it probably wouldn't have been even in their minds, so like … it didn't win, but it helped.

In this respect, activists in Northern Ireland had perhaps a greater appreciation of the potential of court cases than some lawyers involved and worked to ensure that cases went ahead. Of one such case, Goretti Horgan remembered:

> And the fact that she then, that her lawyers, I mean, there was … there was a lot of activist intervention that led to her lawyers realising that there were defences that she could put forward and that it wasn't a case that, you know, she just had to do what everybody else did and plead guilty in order to get it over with. I think that if it hadn't been for the activists' interventions with her barristers etc, that that actually, you know, she may not be taking that case now and of course we crowdfunded to make sure that she was able to get started on the judicial review because there was a big nervousness about it and she … was very frightened about being left with costs and things like that … So … there was activist involvement in ensuring that the case did actually happen.

4.2 International human rights law

As well as domestic litigation, international human rights law emerged as indispensable to the process for change in Northern Ireland and a key location for feminist law work. However, leveraging that body of law required significant legal sophistication, given the constitutional complexities of the devolution settlement. Prior to the devolution of criminal justice powers to Stormont in 2010, activists in Northern Ireland focused their campaigning efforts on Westminster. Following 2010, attention turned to Stormont and lobbying for devolved action. When it became clear that this approach would not be fruitful, activists reframed abortion as a matter of international human rights law compliance, arguing that there was no constitutional impediment to Westminster intervention.

This required a number of critical steps. The first was establishing that the law in Northern Ireland violated human rights and, particularly, the Human Rights Act 1998.

---

57 *A and B* (n 51).

This was achieved through domestic litigation.[58] Sarah Ewart was an intervener in this case. The second was establishing that, as a matter of human rights law, Westminster could be compelled to attend to abortion in Northern Ireland notwithstanding the devolution settlement. This was achieved largely by lobbying MPs and engaging in various parliamentary processes, including the Women and Equalities Committee consultation on abortion in Northern Ireland.[59] Alliance for Choice was central to coordinating and shaping this work. The third was establishing that the law as it stood constituted violations of international human rights law, which was achieved when Alliance for Choice and other organisations[60] successfully requested an inquiry from the CEDAW Committee.[61] In the course of the inquiry, when Committee members visited Northern Ireland, it was found that the law produced manifest violations of international human rights law.[62] This inquiry advanced the (successful) argument that Westminster was not precluded from taking action on abortion and, in fact, that the UK had an international obligation effectively to take steps to resolve those violations of human rights law. The fourth was creating clear imperatives for and guidance in respect of what could be done to resolve these violations at a doctrinal level, including putting an end to ongoing and active prosecution for people participating in (unlawful) self-managed abortion in Northern Ireland. Again, this was significantly achieved by engagement with the CEDAW Committee, which expressly recommended a moratorium on future prosecutions and the cessation of ongoing prosecutions.[63]

As a result of this ability to shift between courts and institutions, from the national to the transnational and back again, activists in Northern Ireland successfully and strategically deployed international human rights law in laying the groundwork for the decriminalisation of abortion in October 2019 and its subsequent regulation. As this suggests, feminist law work requires shifting between – and sometimes working across – different national and international locations simultaneously. Even if activists are focusing on a law-making process outside of parliament, they must also work to ensure that the outcome of that process will be well received by the legislature in future and seek support from international bodies when debate at the national level has become stymied.

In the Republic of Ireland, the dominant and very limiting mainstream interpretations of the 8th Amendment meant there was little opportunity to force the Oireachtas to liberalise abortion law using domestic litigation.[64] Instead, arguments were made in international judicial and quasi-judicial institutions. In particular, although litigation in

58 *Northern Ireland Human Rights Commission* (n 22). In this case violation of Article 8 was held *obiter* following a finding that the Commission lacked standing: Jane Rooney, 'Standing and the Northern Ireland Human Rights Commission' (2019) 82(3) Modern Law Review 535; Bríd Ní Ghráinne and Aisling McMahon, 'Abortion in Northern Ireland and the European Convention on Human Rights: reflections from the UK Supreme Court' (2019) 68(2) International and Comparative Law Quarterly 477.

59 Women and Equalities Committee, Abortion Law in Northern Ireland: 8th Report of Session 2017-19 (HC 1584, 2019).

60 The Family Planning Association for Northern Ireland and the Northern Ireland Women's European Platform.

61 Catherine O'Rourke, 'Advocating abortion rights in Northern Ireland: local and global tensions' (2016) 25(6) Social and Legal Studies 716.

62 CEDAW Committee (n 25).

63 Ibid.

64 For analysis see, for example, Fiona de Londras, 'Constitutionalizing fetal rights: a salutary tale from Ireland' (2015) 22(2) Michigan Journal of Gender and the Law 243.

A, B & C v Ireland[65] did not demonstrate that the 8th Amendment violated the European Convention on Human Rights, it did identify that the absence of any process by which women could assess whether they were entitled to access abortion within the limited constitutional provision constituted a violation of Article 8. As a result, the Oireachtas legislated for access to abortion in 2013 – a full 30 years after the constitutional provision was introduced. While the *A, B & C* litigation failed in its broader objective to establish the rights-incompatibility of the constitutional status quo, it made clear the limitations of the constitutional framework and coincided with the death of Savita Halapannavar, a woman who died following refusal of abortion in Galway in 2012, and whose death ignited widespread activism, bolstering the preceding practices of pro-choice feminist activists.[66] As Ailbhe Smyth put it, 'the case that really did make a difference to our more recent campaigning was the *A, B, C*: the set of three cases that went to the European courts and the European courts gave an instruction to government that they had to clarify the legislative parameters'.[67]

The experience of engaging with the European Court of Human Rights and seeking clear determination that Ireland's law violated human rights nudged activists towards other international institutions, specifically the United Nations (UN). The UN Human Rights Committee found – in *Mellet*[68] and *Whelan*[69] – that the prohibition of abortion in situations of fatal foetal anomalies could constitute a violation of the right to be free from torture, inhuman and degrading treatment or punishment protected in the International Covenant on Civil and Political Rights, further bolstering the arguments for change. While these engagements with international human rights law were useful in the campaign for repeal of the 8th Amendment in Ireland, they were not per se critical to securing the commitments to the Citizens' Assembly and then to the referendum of 2018. Those commitments are better understood as the product of a combination of political coincidence, timing, and relentless and resilient feminist activism.

This engagement with national and international spaces demonstrates our previous comment above that law reform processes are not linear. Recognising often-unexpected spaces, moments and events as landmarks and opportunities to articulate and make claims for reproductive justice and abortion law reform was a critical part of feminist law work, both North and South. Activists had to be ready when the tide turned. National and international challenges surrounding Brexit also offered another opportunity for activists in Northern Ireland. For example, Kellie O'Dowd noted how the minority government in Westminster, reliance on the DUP, Brexit, and resulting tensions between the DUP and other Members of Parliament created opportunities that Alliance for Choice was ready to take advantage of. Indeed, that preparedness to take advantage of such situations was part of their activist strategising:

---

65 *A, B and C v Ireland* Application No 25579/05, Merits, 16 December 2010. For discussion, see Katherine Side, 'A B and C versus Ireland: a new beginning to access legal abortion in the Republic of Ireland?' (2011) 13(3) International Feminist Journal of Politics 390.

66 Máiréad Enright and Fiona de Londras, '"Empty without and empty within": the unworkability of the Eighth Amendment after Savita Halappanavar and Miss Y' (2014) 20(2) Medico-Legal Journal of Ireland 85.

67 The Protection of Life During Pregnancy Act 2013 followed soon after this case and the death of Savita Halappanavar.

68 *Mellet v Ireland* (2016) UN Doc CCPR/C/116/D/2324/2013. For discussion, see de Londras (n 7).

69 *Whelan v Ireland* (2017) UN Doc CCPR/C/119/D/2425/2014. For discussion, see Katarzyna Sękowska-Kozłowska, 'A tough job: recognizing access to abortion as a matter of equality. A commentary on the views of the UN Human Rights Committee in the cases of *Mellet v Ireland* and *Whelan v Ireland*' (2018) 26(45) Reproductive Health Matters 25.

> I suppose … in any crisis you have to take the opportunities that present themselves … the shit show that is Brexit – excuse my language – and the Tory reliance on the DUP has played into our hands because we've had Labour and Conservative MPs who … want to give the DUP a bloody nose … in this area of chaos, we've been able to, I suppose, gain some political capital. And we're going to keep on doing that.

Or as Emma Campbell put it:

> … to misappropriate another famous Irish phrase, England's misfortune is Northern Ireland's opportunity at the minute and we're trying to make sure that we're not going to leave them alone about abortion in Northern Ireland and we really hope that it causes them a massive pain in their side.

This key demand of feminist law work – to be active at many sites at once and to 'get our fingers in all the abortion pies' (Emma Campbell) – played out differently on each side of the Irish border. Given Northern Ireland's place within the post-colonial, and soon-to-be post-Brexit, UK, Northern Ireland activists worked not only in varied locations of legislatures, courts and international processes, but also in various jurisdictions. They undertook feminist law work in Northern Ireland, in Westminster, and in the Republic of Ireland where they substantially supported efforts for Repeal, both as a marker of the all-island nature of abortion rights activism and because success in the South held significant potential for momentum in the North.

4.3 Legal education in the public sphere

The wider public sphere is also, of course, a site of activist legal mobilisation. As McCann has written of social movement litigation, legal activism has 'radiating' or 'spillover' effects and value beyond the formal legal process and its instrumental gains and losses, through moments of engagement with legal institutions and opportunities increasing public rights consciousness, for instance.[70] Much feminist law work goes into translating[71] or 'vernacularising'[72] new legal claims and norms wrung from legal processes for everyday usage and understanding.[73] This activity was evident in the abortion rights campaigning on the island of Ireland as well. Erin D'Arcy describes in detail how she used 'In Her Shoes' to translate legal concepts for ordinary members of the public engaging with her website:

> Maybe a year ago or so I probably wouldn't have been able to pronounce [the word] autonomy or know what it meant. So, I had to look it up and go, 'What does autonomy even mean? I've heard this word, I've seen it, I don't even know how to say it'. And I feel like a lot of people have no idea and so these words are all being used by professionals and people who are, you know, knowledgeable or whatever and it really alienates a lot of people and so they feel like, 'I don't know what this is, I don't know how to say it'. And so when I realised that myself … I kind of tried to help break that down a little bit … So like I would use the word autonomy and then I would also include 'the right to govern myself' in the explanation. So that way people who are reading ... would start to piece that together

70 McCann (n 35).

71 Ruth Fletcher '#RepealedThe8th: translating travesty, global conversation, and the Irish abortion referendum' (2018) 26 Feminist Legal Studies 233.

72 Sally Engle Merry, *Human Rights and Gender Violence: Translating International Law into Local Justice* (University of Chicago Press 2006).

73 For wider work on this point, see Marie-Lise Drapeau-Bisson, 'Beyond Green and Orange: Alliance for Choice-Derry's mobilisation for the decriminalisation of abortion' (2020) 35(1) Irish Political Studies 90.

Emma Campbell also described an important part of Alliance for Choice's current strategy as being focused on 'educating people mostly in England about what Northern Ireland is and the fact that we don't have an abortion law.'

Accordingly, activists are simultaneously law-makers and strategic teachers of law. They not only help the public to understand what law is but also make them aware of opportunities to engage in the democratic process around law and law reform. The activist role in translating the outcomes of formal legal processes for a wider political audience is crucial to legal change. However, legal discourse is not necessarily 'made' at one site and 'translated' at another; feminist law work means active intervention in both spaces and is directly generative of new legal norms. Activist legal agency is not confined to reacting to missives from 'a distant, official terrain wholly defined by elite judges, lawyers and professional legal commentators'.[74] As this account of feminist law work shows, that terrain is neither distant nor necessarily official, and is formed by and through imaginative and insistent activist interventions.

5 The 'everydayness' of feminist law work

Having pieced together a view of the multiple spaces and complex approach to law in feminist law work, it is necessary to turn to the everyday nature of this work. What are the everyday practices and activities that stimulate legal change? The discussions emerging from our interviews and workshop indicate that there are a number of responses to this question. We will focus on three: adaptability and skill; narrative production and management; and labour.

5.1 Adaptability and developing new skills

First of all, and perhaps definitively, feminist law work requires feminist activists to be adaptable. Such work often demands the development of new skills or knowledge to respond to shifting constitutional, legal, political or social opportunities and developments.[75] This is particularly evident in the context of Northern Ireland. Following the collapse of the devolved Assembly at Stormont in January 2017 the jurisdiction did not have a functioning devolved government until January 2020. This meant that there was not a working legislative institution in Northern Ireland through which formal law reform could be advanced. There was also a clear acceptance that, even if it were functioning, Stormont was a hostile place for abortion law reform. As Kellie O'Dowd explained in a pre-workshop interview:

> We have fucking research after research, we have just a legislature that are not fucking interested, so we need to take this opportunity … [I]f they legislated, if they were pushed, through public opinion, to legislate it would be FFA [fatal foetal anomaly] and maybe sex crimes, they might even wiggle out of that, and it would be tiny legislative change, so the hope is that we get this sorted out before they get back.[76]

74 Michael McCann and George I Lovell, 'Towards a radical politics of rights' in Paul Christopher Grey (ed), *From the Streets to the State: Changing the World by Taking Power* (SUNY Press 2010) 139, 141. See also Michael McCann, 'Reform litigation on trial' (1992) 17(4) Law and Social Inquiry 715, 731

75 Emma Campbell, 'From grassroots to government: arts engagement strategies in abortion access activism in Ireland' in Colleen MacQuarrie, Fiona Bloomer and Claire Pierson (eds), *Crossing Troubled Waters: Abortion in Ireland, Northern Ireland, and Prince Edward Island* (Island Studies Press 2018).

76 Goretti Horgan, by contrast, wondered whether, if abortion was decriminalised by Westminster, Stormont would be capable of legislating on abortion at all.

As a result – as noted above – feminist groups turned their focus to Westminster, to primary legislation, and to issues of constitutional law and devolution, all in the context of ongoing negotiations regarding the UK's planned exit from the EU. In navigating this shift, feminist activists not only had to develop a host of new arguments (about devolution, for example, and the question of whether abortion is a reserved or devolved issue), but also adapt to new political challenges in a context where the predominant sentiment in the Westminster Parliament is at least somewhat pro-choice. Emma Campbell reflected on what this has required of activists:

> There's a slight issue with the Northern Ireland campaigning at the minute having to be at Westminster … we are not convincing MPs in Westminster to be pro-choice. The house is pro-choice. We are talking about issues of devolution, we're talking about the Sewell Convention … I've read the Good Friday Agreement for no thanks whatsoever.

In this respect, activists are required to respond to political and legal changes, look to different areas of law, and expand their networks and resources in these areas. They must do so often in very short time-frames, needing to be ready to offer quick, digestible analysis, proposals, expertise, and arguments to a range of audiences, even in highly technical and sometimes unclear or uncertain terrains of law.

This adaptability means that we can understand legal claims-making not only as a tool which facilitates campaigning but also as a resource that shapes and changes social movements.[77] As activists have become experienced in law reform processes – albeit approaching law often rather instrumentally – they have invested greater resources and energies in legal mobilisation, leading to shifts in future tactics and in their self-understanding as legal actors. The challenge that this poses is exacerbated when we consider that, like many feminist activists, they must adapt, not only to single mode of feminist law work (formal legal change) but also between modes of work. In particular, they must adapt relationships between policy and law reform work and direct assistance-based activism for abortion-seekers.[78] In addition to being lay-lawyers and legal commentators, our participants are also involved in activity to support abortion-seeking women, to help them practically, to offer counselling, to engage in training and awareness-raising campaigns, to engage with the media on a regular basis, and to construct a voice on social media. These are disparate forms of activity requiring differing skills, a degree of self-training and an ability to switch between these positions on a daily, weekly or hourly basis. As Susan Buck-Morss has observed, when activists teach and train themselves, they also empower themselves for future action.[79] In time, Alliance for Choice found that political and legal allies came to rely on them so that being able to fulfil that role becomes central to their ability to influence the legal process. As Kellie O'Dowd says:

> In the morning if [abortion was decriminalised] and there had to be a submission on well what do you want the new law to look like, I have no doubt that politicians would call me. You know I have no doubt about that. Because they haven't a fucking clue and they know we're the real experts.

77 Colm Campbell and Ita Connolly, 'A deadly complexity: law, social movements and political violence' (2007) 16 Minnesota Journal of International Law 265.

78 Judith Orr, *Abortion Wars: The Fight for Reproductive Rights* (Policy Press 2017) 25–27.

79 Susan Buck-Morss, 'Democracy: an unfinished project' (2014) Boundary 92.

5.2 Narrative production and management

Notably, feminist activists are involved in narrative production, and management of these narratives.[80] Social movements literature emphasises the role of activists in 'framing' social problems in ways that convince sympathetic members of the public of the need for collective action.[81] A frame is a simplified, condensed reinterpretation of a social problem. It often reworks older, culturally embedded interpretations, amplifying some aspects and extending others. Perhaps one of the most important elements involved in feminist law work on abortion, then, is the framing, production, curation and advancement of narratives from and about women's, and others', experience of current legal and healthcare frameworks regulating abortion.[82] These narrative accounts have proven powerful and affective in both jurisdictions as a way to educate about abortion and to stimulate public discussion; they make concrete the real life of abortion law and can underpin significant shifts in public opinion (held and/or expressed). This is evident, for example, in the public mobilisation and shifts in political discourse that occurred following Savita Halappanavar's death in Galway in 2012 and Sarah Ewart's decision to speak of her experience of being denied abortion care in Northern Ireland in 2013. For our participants, such narratives were a way of making law real and affective, showing its impacts and calling for accountability. They recognised the power of this narrative production, as well as its costs for those whose lives and experiences were publicised, as made clear in the exchange between Kellie O'Dowd and Sarah Ewart already recounted above.

These narratives emerge as a means through which feminists can seek to 'hack' the legal system. By this we mean that they can stimulate a change in debate when other sources, including law, cannot; such narratives can be employed to break deadlock or stagnation in discussion on abortion law and its everyday effects. In particular, as Francesca Polletta has argued, narratives rely on 'emotional identification or familiar plots rather than on testing or adjudication of truth claims'.[83] Stories of personal experience are difficult for opponents to negate. In addition, they give abstract legal frames a plot and map them onto a set of lived experiences.[84] For this reason, narratives can connect people to claims for justice in the way that 'simple' legal or political arguments might not; as Kellie O'Dowd said:

> [H]uman rights doesn't work. Bodily autonomy doesn't work. Feminist arguments don't work. Your story, real life stories, do absolutely work.

80 Lisa Hallgarten, 'Abortion narratives: moving from statistics to stories' (2018) 391(10134) Lancet (2018) 1988.

81 David Snow and Robert Benford, 'Master frames and cycles of protest' in Aldon Morris and Carol McClurg Mueller (eds), *Frontiers in Social Movement Theory* (Yale University Press 1992); Dawn McCaffrey and Jennifer Keys, 'Competitive framing processes in the abortion debate: polarization-vilification, frame saving, and fame debunking' (2000) 41(1) Sociological Quarterly 41; Judith Butler, *Frames of War: When is Life Grievable?* (Verso 2016).

82 These often compete with other frames, such as those produced by the media: Orla McDonnell and Padraig Murphy, 'Mediating abortion politics in Ireland: media framing of the death of Savita Halappanavar' (2019) 16(1) Critical Discourse Studies 1.

83 Francesca Polletta, '"It was like a fever …" narrative and identity in social protest' (1998) 45(2) Social Problems 137.

84 Kristine Olsen, 'Telling our stories: narrative and framing in the movement for same-sex marriage' (2014) 13 Social Movement Studies 248.

At the same time, personal narratives afford a subject position to those telling their stories that formal legal process has not offered.[85] As Sarah Ewart explained:

> Yeah, in court you're very much just a case. You're not really a person. I find when like when we met with [a conservative public figure], she said 'oh I read your stuff in the papers'. But when she came to the house and actually heard what all had happened, she was like 'I didn't realise'. But she sat in tears at different parts of me telling her things. So, I think telling it like this and meeting with people is completely different to how it's come through court ...

An important part of feminist law work in this area, therefore, is framing these stories and making decisions on how they can best be used, while maintaining respect for the position of the woman who speaks. This was a prominent theme in Erin D'Arcy's account of managing 'In Her Shoes'. Erin reflected on her role as one of careful, political curation:

> I didn't have like rules to follow of what stories I could tell and what stories I should not tell but I knew that there was a responsibility in what stories were going out and the amount of people that were reading them and how it would impact their vote … there were some stories that … I knew wouldn't be received well given what had been going on in the media at the time and so I … had to hold it until maybe after the referendum or a different day, depending on what was going on.

Through its effective production of narrative, 'In Her Shoes' offered people a way to have difficult conversations, on and offline, and had the potential to appeal to a wide variety of people who were often outside the traditional remit of feminist activism. Speaking of the website, Linda Kavanagh said:

> … it was so, so important and for women to be able to tell their stories anonymously and then for people to be able to share it in a way that was non-confrontational was really, really powerful, and it really, really worked.

However, narrative production is also a weighty and a challenging part of feminist law work. Our participants demonstrated keen awareness of the ethical concerns that come with narrative production and management. These include the privacy of the individuals involved, including some caught up with law, who perhaps did not want to have their stories told in such a public way. Feminist law work's commitment to women's agency, autonomy and lived experience creates ethical compulsion towards taking care in the production of these narratives. Erin D'Arcy's story of the daily labour involved in curation of 'In Her Shoes' illustrates this in clear terms. By the end of the referendum campaign, 10–15 people a day were sending their abortion stories to her:

> I kind of felt a lot of personal responsibility that when they were sending me these stories that I needed to be there because I was very, very aware that people were anxious to send something like that, and also that they were really vulnerable sending that. A lot of times women would tell me that it was their first time ever telling somebody or their first time every writing it down and how healing that was for them. So, I did feel a lot of responsibility to make sure that I was there. And being there also meant that I had to monitor all the comments coming in so … I kind of had to be actively there on my phone 24-7 really, which meant that at 2 o'clock in the morning when I woke up to nurse my baby, that I'm going through my phone to check the notifications to check comments and

85 On feminist critiques of courtroom processes as objectifying, see e.g. Sara Cobb, 'Transcribing the body and materializing the subject' in Michael Hupek and Gary P Radford (eds), *Transgressing Discourses* (SUNY Press 1997) 195.

> make sure that they're okay and to make sure that the page is okay and nothing's been attacked, it hasn't been – nothing you know happened to it. Because [until I drafted some friends in to help me] it was just me doing it and I really didn't want any person who shared their story to wake up in the morning and see horrible comments. I wanted to protect them from that because … they shared so vulnerably and beautifully with me. So yeah, I lived on my phone constantly.

Despite the work done to care for those who speak, there is, of course a feminist dilemma here, not only for the use of stories that people never intended to tell, but also in the idea that narratives and stories must be retold as a condition of women being listened to. In this process there is a risk that women become exploited, commodified and sacrificed. Erin D'Arcy was aware of this:

> [W]e shouldn't have to use their stories only to be able to get a law for ourselves … I don't know if law-makers are understanding that, that we shouldn't have to have Savita [Halappanavar's] name and her face on something for her to be able to get the care she needs. We shouldn't even know her name.

Negotiating effectively with law-making institutions involves certain compromises.[86] As Reva Siegel has argued,[87] in order to influence legal change social movements may be required to conform to the 'public value condition', framing their demands in a discourse that resonates with notionally shared cultural norms. This is a political move, which is inevitably shaped by dominant power relations in society.[88] In their narrative production work activists may also internalise, or at least bargain with, some of the bureaucratic or procedural constraints of legal institutions or media,[89] with important consequences for advocacy. As Ailbhe Smyth commented:

> [O]bviously there is a selection process and that in itself also raises ethical questions. Of course it does. Because you are either explicitly or implicitly thinking about stories which will actually carry the kind of impact and weight and meaning and emotion and message that you want them to carry.

Curating narratives in this way has important disadvantages; as Davina Cooper notes, it may silence non-hegemonic voices and risk conflict and division which will require labour to avoid or repair.[90] The risks and problematic practices involved must be weighed up by the feminist activist alongside the desire to maintain a feminist perspective and voice in public debate, especially in the media. As Ailbhe Smyth says of the media, 'if you don't give them something, they will go and get it somewhere else'. These are difficult judgements in feminist law work, and these appear as a central part of managing narrative as an important but risk-infused tool.

---

86 Alba Ruibal, 'Social movements and constitutional politics in Latin America: reconfiguring alliances, framings and legal opportunities in the judicialisation of abortion rights in Brazil' (2015) 10(4) Journal of the Academy of Social Sciences 375.

87 Reva Siegel, 'The jurisgenerative role of social movements in United States constitutional law' <https://law.yale.edu/sites/default/files/documents/pdf/Faculty/Siegel_Jurisgenerative_Role_of_Social_Movements.pdf>.

88 Myra Marx Ferree, 'Resonance and radicalism: feminist framing in the abortion debates of the United States and Germany' (2003) 109(2) American Journal of Sociology 304.

89 Anne Revillard, 'Social movements and the politics of bureaucratic rights enforcement: insights from the allocation of disability rights in France'(2017) 42(2) Law and Social Inquiry 450.

90 Davina Cooper, 'An engaged state: sexuality, governance, and the potential for change' (1993) 20(3) Journal of Law and Society 257.

5.3 Labour

Feminist law work, of course, involves labour, little of which is remunerated. Persistence and a commitment to long-term activity towards change at individual and collective levels are essential for feminist activists.[91] Sometimes the labour that is undertaken is a slow chipping away over time. In the Northern Irish context Sarah Ewart's comments demonstrate an awareness of and a commitment to this kind of persistence: 'we'll just keep going until we try and get it changed. Five years later.' At other times, however, labour is not so demanding in terms of its duration but, rather, in its intensity. This was particularly evident from those involved in the referendum campaign in Ireland. Linda Kavanagh reflects on how this period stimulated a disorientating feeling for activists in hindsight:

> [I]t's really obviously very nice to be on the other side of that … it wasn't that long ago although time has no meaning for me anymore and it feels like decades.

The labour involved in feminist law work, furthermore, must be understood as evolving over time. This evolution must be carefully reflected on and responded to strategically. For example, at the time of our workshop the challenge for those in Ireland was no longer about mobilising support and canvassing for votes, but about motivating supporters to continue in work to secure satisfactory legislative and healthcare frameworks for abortion access on the ground. This required a shift in focus not only for leaders, but for all in the movement, and there are attendant risks of premature demobilisation. As Linda says, this kind of work is 'not as sexy as getting out on the street. It's hard to keep people motivated.'

Throughout our discussions, the emotional toll of this labour, and the demands made on activists by its various forms and shifts were clear. This is not just physical labour, although that is of course involved too, but emotional labour.[92] The idea of 'feminist law work' has the potential to bring into view the emotional life of doing and engaging with law as non-lawyers,[93] a life that is often rendered invisible when we approach law and its reform through more traditional lenses. Both Goretti Horgan and Linda Kavanagh recalled incidents where they were moved to tears by the emotional toll of the unpredictable reform effort – a sudden victory or loss – or the stories of women which they were confronted with which stimulated an affective and empathetic response. Linda Kavanagh describes her experience:

> I was crying, you know, at my desk because another story had broken, like Miss Y had broken or the woman in the coma[94] and … feeling like it was never going anywhere.

Activists also expressed a variety of other emotions attached to their labour, including dissatisfaction and frustration. As Emma Campbell says of Alliance for Choice, 'we're just used to always being dissatisfied'. This distinct emotional experience can set activists apart

91 Cynthia Enloe, *The Big Push: Exposing and Challenging the Persistence of Patriarchy* (University of California Press 2017).

92 Jeff Goodwin, James Jasper and Francesca Poletta (eds), *Passionate Politics: Emotions and Social Movements* (University of Chicago Press 2001); Gavin Brown and Jenny Pickerill, 'Space for emotion in the spaces of activism' (2009) 2(1) Emotion, Space and Society 24; Suzanna Franzway, 'Women working in a greedy institution: commitment and emotional labour in the union movement' (2000) 7(4) Gender, Work and Organization 258.

93 For discussion on emotional labour of actors working within law and legal institutions, see John Hagan and Fiona Kay, 'The emotional toll and exhilaration of human rights activism: gender and legal work at the Hague International Criminal Tribunal' (2011) 37(1) Queen's Law Journal 257.

94 *PP v HSE* [2014] IEHC 622.

from their wider families and communities. Linda Kavanagh remembered an illustrative incident in the period after the Repeal referendum victory:

> Actually, funnily enough, I was at a funeral a while ago and somebody said … we ended up talking about the referendum and somebody was kind of asking what I was doing now. And I was like, I'm taking it a bit easy, it took a lot of out me. And they're like, 'Yeah, 64% [the total Yes vote in the referendum] must be really, really difficult to deal with'. Like really sarcastically. And I was just like, oh, you don't get it. At all. So, the layperson, that is how they will perceive it. They will perceive it as it was nice. It was tough but it was nice in the end, so we're all fine.

For those who have been personally affected by the law, there is a particular emotion involved in their labour which must be recognised. Jane Christie explains the difference simply: 'We're living it. And I think that impacts more.'

While our participants clearly evidenced the demanding and emotional nature of their labour, they also spoke of what sustains them in this work. Often the work is enabled by the relational bonds of their movements,[95] and the small moments of progress that they can detect. Again, Linda Kavanagh captured this:

> I suppose the thing possibly that keeps you doing is your community, you know, the other activists, and then sometimes it's really weird, the small things … you kind of sometimes take your victories where you can … you kind of have to take them where you can, because it can seem so long and it can be so dark … you don't get a lot of May 26ths.

Conclusion

In this account we have drawn attention to and named feminist legal work as an important form of activity at the heart of legal change processes; one that is often under-considered or rendered invisible by dominant reform narratives. Feminist activists are an often overlooked or invisible part of the plural, non-linear and contingent journeys towards abortion law reform in both Ireland and Northern Ireland. They are not the whole story of reform, but their work is an important part of it. In drawing attention to these non-legal actors, we have not sought conclusively to define, delineate or deconstruct feminist law work, but to offer a sense – from our engagement with eight activists – of the variety, challenges, skills and indispensability of feminist law work to formal processes of abortion law reform in Ireland and Northern Ireland. Thinking about this manifold, instrumental, creative, emotional and laborious activity as work which is equally important as that undertaken by legislators, lawmakers, judges, and other legal arbitrators offers an additional thread to reflections on the Northern/Irish journeys towards legal change in this area.

More than just a way of making sense of how commitments to change the law were achieved, drawing attention to this work and naming it also holds potential to assist us in understanding the current legal landscape and ongoing activity on abortion law in both jurisdictions. The feminist law work continues even as formal law reform is achieved: it is an ongoing part of shaping new legislation, of publicising legal failure, of highlighting the gaps between law and practice, of insisting on a richer conception of reproductive justice, and in subverting, playing at the edges of, 'being cheeky'[96] towards, and 'tickling' (Emma Campbell) law.

---

95 Mario Diani and Doug McAdam (eds), *Social Movements and Networks: Relational Approaches to Collective Action* (Oxford University Press 2003).

96 Fletcher (n 20).

For us as feminist legal scholars, the activists who undertake this work are a vital reminder that it is possible, and necessary, to reimagine law and what is involved in formal legal change processes. They encourage us to consider again (and again) how we understand the limits and excesses of law and the forms of legal knowledge and skills that we value, and to insist on inclusive understandings of what it means to engage with and 'do' law in everyday life.

NILQ 71(3): 387–400

The 'one man company' after *Patel v Mirza*: attribution and illegality in *Singularis Holdings v Daiwa Capital Markets*

JAMES C FISHER

Sophia University, Tokyo, Japan

Abstract

This note discusses the UK Supreme Court's decision in Singularis Holdings v Daiwa Capital Markets *in the context of other recent decisions on corporate attribution and the illegality principle in English law. It particularly considers* Daiwa*'s implications for the relationship between the illegality doctrine and other legal principles in the wake of* Patel v Mirza. *The court employed a context-sensitive, teleological approach to attribution, one consequence of which was the conclusive consignment of the House of Lords' decision in* Stone & Rolls Ltd v Moore Stephens *to irrelevance. It nonetheless privileges orthodox, pre-*Patel*ian authority in the disposal of the case. The court's approach suggests that* Patel *is perceived as the high-water mark for expansive, policy-sensitive understanding of the illegality principle, and that its disruptive potential is likely to be carefully constrained in future decisions of the Supreme Court.*

Keywords: *Singularis Holdings v Daiwa Capital Markets;* illegality; attribution; *Patel v Mirza*; *Stone & Rolls Ltd v Moore Stephens*; *ex turpi causa*; directors' duties; fraud.

Introduction

The facts of *Singularis Holdings v Daiwa Capital Markets*[1] were simple and not long ago would hardly have warranted the attention of the Supreme Court. Nonetheless, the recent volatility in the English illegality doctrine forced from the court an authoritative restatement of the law concerning the corporate attribution of a director's unlawful conduct.

Violent disagreement about the illegality principle's nature, justification and operation is inevitable since, to a degree far greater than most questions of legal doctrine, the principle forces lawyers to confront some of the most profound theoretical and ideological questions that divide juristic thinkers. It raises uncomfortable questions about the distinction between legal and non-legal arguments, and therefore about law's normative autonomy as a site of structured, internally self-validating reasoning. It moreover exposes the tension between (i) the equality, clarity and predictability promised by uniform legal rules and (ii) episodic justice in concrete disputes, which is maximised by flexibility and judicial discretion.

1 [2019] UKSC 50.

For a short while, it seemed as though the Supreme Court's 2016 decision in *Patel v Mirza*[2] had settled – at least as a matter of positive law – the debate over the proper operation of the illegality principle, which had received a flurry of judicial attention at the highest level since *Hounga v Allen*[3] was decided in 2014. However, just as the law seemed conclusively to have rejected a formalistic, procedural rule in favour of a flexible principle attentive to public policy, a sequence of Court of Appeal decisions resisted litigants' suggestions that *Patel* invited departure from older authorities. Appellate decisions in the wake of *Patel* have generally remained loyal to older cases which sit imperfectly with *Patel*'s approach to illegality. *Daiwa* continues this trend, for the first time at the Supreme Court level.

1 Facts, arguments and disposal

Singularis Holdings was a Cayman Islands-incorporated company that existed to manage the personal wealth of Saudi Arabian tycoon Maan Al Sanea. He was the sole shareholder as well as the chairman, president and treasurer. Although there were directors other than Al Sanea, he had been delegated sweeping management powers, particularly with respect to authorising instructions to the company's bankers. When the company's financial situation became fragile, Al Sanea instructed its bankers, Daiwa Capital Markets, to make a series of payments from the company's account to two other companies – a misapplication of company funds in breach of Al Sanea's fiduciary duties.[4]

Singularis subsequently entered voluntary liquidation and its liquidators sued Daiwa for recovery of the sums paid away according to Al Sanea's instructions,[5] on the basis either (i) that Daiwa had dishonestly assisted in Al Sanea's breach of his fiduciary duty, or (ii) that Daiwa had breached the duty of care recognised in *Barclays Bank plc v Quincecare Ltd*.[6] That duty obliges banks to investigate before executing orders from a director of a client company if the circumstances lead reasonably to suspicion that such orders constitute a fraud against the company. A bank is liable if it executes an order knowing it was made dishonestly, ignoring obvious dishonesty, or recklessly having failed to make the confirmatory enquiries a reasonable, honest person would make in the circumstances.

Rose J dismissed the first argument, holding that Daiwa had acted honestly and was therefore not liable for assisting in the misapplication of company funds. The dishonest assistance point was not appealed, so the Court of Appeal and Supreme Court dealt only with Daiwa's liability in negligence for breach of its *Quincecare* duty of care. At trial, Rose J held that Daiwa had indeed breached its duty to Singularis, identifying 'many obvious, even glaring, signs that Mr Al Sanea was perpetrating a fraud on the company'.[7] She awarded the sums claimed, less a 25 per cent reduction reflecting the contributory negligence attributable to Singularis. The Court of Appeal unanimously upheld her decision.[8]

2 [2016] UKSC 42, [2017] AC 467.

3 [2014] UKSC 47, [2014] 1 WLR 2889.

4 [2017] EWHC 257 (Ch), [2017] Bus LR 1386, [120]–[127].

5 Less whatever was recovered from Mr Al Sanea personally or from the recipients of the impugned payments.

6 [1992] 4 All ER 363.

7 *Daiwa* (n 4) [192], per Rose J (further specifics at [193]–[202]).

8 [2018] EWCA Civ 84, [2018] 1 WLR 2777.

The Supreme Court thought it 'incontrovertible' that Daiwa had breached its *Quincecare* duty to Singularis,[9] and that Rose J's decision must stand unless there existed some impediment to the negligence action. Daiwa submitted that Singularis' otherwise valid negligence claim must nonetheless fail for one (or more) of three reasons. Each of these putative obstacles to Singularis' claim presupposed that Al Sanea's fraud against the company should be attributed to the company itself.

First, Daiwa argued that Singularis' claim was incompletely constituted for want of causation since its losses were legally attributable not to Daiwa's negligence, but to Singularis' own. Second, and connectedly, Daiwa argued that Singularis' claim in negligence was defeated by Daiwa's own 'equal and opposite' claim in the tort of deceit.[10] Third, Daiwa argued that Singularis' claim was barred by the illegality rule as formulated in *Patel*. In a concise single judgment, the Supreme Court upheld the decisions of Rose J and the Court of Appeal, declaring that 'for the purpose of the *Quincecare* duty of care, the fraud of Mr Al Sanea is not to be attributed to the company' and that, even if it were so attributed, 'none of the defences advanced by Daiwa would succeed'.[11]

2 Causation and/or a countervailing action in deceit

Daiwa invoked *Reeves v Commissioner of Police of the Metropolis*[12] to argue that Singularis' negligence claim must fail on causal grounds. In *Reeves*, Lord Hoffmann had identified 'a difference between protecting people against harm caused to them by third parties and protecting them against harm which they inflict upon themselves'.[13] Daiwa submitted that, through the actions of its fraudulent director, Singularis had in law inflicted the losses in question on itself, and they were therefore not causally attributable to Daiwa's breach of duty. Rather, this was a case in which the claimant 'must look after themselves and take responsibility for their actions'.[14]

The Supreme Court swiftly exposed the flaw in that submission. The distinction that Lord Hoffmann identified relates to the judicial decision whether or not to recognise a novel duty of care. For reasons of policy and ideology, the law will be far slower to recognise a duty to prevent another person harming their own interests than a duty to protect against harm inflicted by third parties. The Supreme Court in *Daiwa* approved Lord Hoffmann's explanation that, in determining whether such an exceptionally recognised duty has been breached, it would be 'self-contradictory to say that the breach could not have been a cause of the harm because the victim caused it to himself'.[15]

The crucial, and sensible, point in Lord Hoffman's opinion is that the law should explicitly accept the dilemma surrounding liability for others' self-inflicted harm as *normative*, rather than disguise it as an analytical problem about causality. The clarity and integrity of the law both benefit from such transparency. Obviously, in these situations both the claimant's action and the defendant's omission to prevent that action (or its consequences) are historically necessary conditions – 'but-for causes' – of the claimant's loss. The real dilemma concerns what allocation of responsibility between these causal agents is just and appropriate.

9 *Daiwa* (n 1) [12].
10 Ibid [1].
11 Ibid [38].
12 [2000] 1 AC 360.
13 Ibid 368.
14 Ibid.
15 Ibid.

In *Daiwa*, the Supreme Court correctly explained that this dilemma had already been authoritatively settled, as reflected by the emergence of *Quincecare* liability in the first place. The *Quincecare* duty represents a balance between competing interests and policies in obliging banks to prevent their customers suffering losses necessarily 'caused by people for whom the customer is, one way or another, responsible'.[16] It follows that liability for breach of that duty could not be avoided by the causal analysis Daiwa deployed.

The Supreme Court employed the same analysis to dispense with the submission that Daiwa's countervailing action in deceit meant that Singularis' negligence claim must be dismissed for circularity. It adopted the insight of the Court of Appeal, which had found 'surprising' the suggestion that Daiwa could escape liability for breach of its *Quincecare* duty – a duty to act against precisely this kind of fraud – by invoking 'the fraud that was itself a pre-condition for its liability'.[17] The court's dismissal of these ambitious defensive submissions maintains the coherence of the law in this area. *Quincecare* liability can, after all, only arise in the presence of this kind of fraud and is hardly comprehensible as a doctrine if the presence of such fraud defeats the liability that depends on it.

3 Illegality

Both Rose J and the Court of Appeal also rejected Daiwa's submission that Singularis' action was barred by the illegality rule, even assuming that Al Sanea's actions were regarded at law as the actions of his company.

Patel, now the leading case on the illegality principle, requires courts to determine the engagement or otherwise of the illegality principle by considering:

> (a) … the underlying purpose of the prohibition which has been transgressed and whether that purpose will be enhanced by denial of the claim, (b) … any other relevant public policy on which the denial of the claim may have an impact and (c) … whether denial of the claim would be a proportionate response to the illegality, bearing in mind that punishment is a matter for the criminal courts.[18]

In light of this authoritative formula, argument before the Supreme Court in *Daiwa* concerned: (i) whether allowing or barring the claim would better promote the integrity of the legal system; (ii) the contribution of *Quincecare* liability to the important goal of countering financial crime; and (iii) the proportionality of the illegality defence in light of the wrongdoing attributable to Singularis.[19]

Ultimately, the Supreme Court decided the case on attribution grounds, declining to attribute Al Sanea's wrongdoing to Singularis and thus excluding all Daiwa's defences, including the illegality principle. Nonetheless, the court examined and approved the inferior courts' approach to illegality.[20]

Rose J had reasoned that for the illegality principle to obstruct claims such as this would have negative public policy implications. Specifically, reducing the consequences for banks who negligently fail to scrutinise potentially fraudulent actions would endanger banks' role in combatting financial crime.[21] Moreover, 'denial of the claim would be an unfair and disproportionate response to any wrongdoing on the part of Singularis'.[22] The

16 *Daiwa* (n 1) [23].
17 *Daiwa* (n 8) [79]; endorsed by the Supreme Court (n 1) [25].
18 *Patel* (n 2) [120], per Lord Toulson.
19 Summarised by the Supreme Court (n 1) [20].
20 Ibid [21].
21 *Daiwa* (n 4) [219].
22 *Daiwa* (n 1) [18].

Court of Appeal upheld this approach and reiterated the need to preserve the substance of the *Quincecare* duty, which served to promote legality and accountability in finance. The total failure of Singularis' claim – the necessary consequence of a successful illegality defence – would be disproportionate because Daiwa's negligence was particularly egregious.[23] For the Court of Appeal, greater fairness would be achieved by allowing the claim to succeed but reducing Singularis' damages to reflect its own contributory negligence.

The Supreme Court recalled this reasoning without any suggestion that the lower courts' understanding or implementation of the illegality doctrine was flawed. While Lady Hale (for the court) reiterated Lord Neuberger's declaration in *Patel* that assessing the effect of the illegality principle on a claim is 'not akin to the exercise of discretion',[24] she did not suggest that the lower courts' approach had offended this.

4 The logically prior question: attribution

All Daiwa's oppositions to liability for breach depended on the central contention that, because Al Sanea was the sole 'controlling mind and will' of the company,[25] his wrongdoing must be regarded at law as the company's own, a proposition that both Rose J and the Court of Appeal had rejected.[26] The Supreme Court sought to go 'back to basic principles' to decide the issue of attribution and, ultimately, the relevance of illegality. This may be read as tacit acceptance that the question of attribution in the context of illegality had become muddied as a result of rapidly changing understandings of the illegality principle and its relationship with other legal doctrines.

The court's 'starting point' was the central axiom of company law; the distinct legal personality of properly incorporated companies.[27] Fidelity to that principle doomed a rather desperate final defence attempted by Daiwa, namely that 'the law should not treat a company more favourably than an individual'.[28] To this, the Supreme Court retorted simply that 'companies are different from individuals', precisely because of their distinct legal personality[29] and the consequent need for rules of attribution to determine when human actions are to be regarded in law as the actions of that distinct legal person. A bold argument of principle that conducting business as a company (rather than as private persons) should not affect the legal liabilities of the human beings engaged in enterprise was never likely to be favourably received. After all, to alter that legal position *vis à vis* the world at large is the whole point of incorporation. The English courts are abidingly reticent to articulate systematically the outer limits of corporate personality, and 'veil piercing' consequently remains under-theorised.[30] Although the law may depart from the

23 *Daiwa* (n 8) [66].

24 *Patel* (n 2) [175].

25 *Daiwa* (n 1) [26].

26 *Daiwa* (n 4) [208]–[215] and (n 8) [50]–[60].

27 *Salomon v A Salomon and Co Ltd* [1897] AC 22.

28 *Daiwa* (n 1) [37]. Daiwa had cited for comparison *Luscombe v Roberts* (1962) 106 SJ 373, in which a solicitor failed against his negligent accountants because he knew that his own actions were unlawful.

29 *Daiwa* (n 1)[34].

30 Pey Woan Lee, 'The enigma of veil piercing' (2015) 26(1) International Company and Commercial Law Review 28; Marc Moore, 'A temple built on faulty foundations: piercing the corporate veil and the legacy of *Salomon v Salomon*' (2006) JBL 180. Cf the more frequent veil-piercing in jurisdictions with a more socially integrated approach to the limits of the corporate form: Kimberly Bin Yu and Richard Krever, 'The high frequency of piercing the corporate veil in China' (2015) 23(2) Asia Pacific Law Review 63.

'unyielding rock' of *Salomon v Salomon*,[31] the power to do so is deployed with extreme trepidation.[32] The company's separate personality – 'the whole foundation of English company and insolvency law'[33] – will not be disregarded just because it impedes some other legal policy or initiative.[34] Separate personality does – and *must* – frustrate certain other legal priorities if the corporate form is to be meaningful. Certainly, there have always been forceful normative objections to the corporate form's capacity to absorb or dissipate liabilities that would otherwise settle on identifiable persons of flesh and blood. But such grand first-order debates of legal and social policy are primordial, overtly political and beyond realistic challenge in a specific attribution dispute.

From this starting point, the Supreme Court proceeded to describe how the law must determine when acts of one person (a director) are treated as those of another (the company). It invoked the 'classic exposition' of this investigation,[35] namely Lord Hoffmann's speech in *Meridian Global Funds Management Asia Ltd v Securities Commission*.[36] In characteristically Hoffmannian fashion, his Lordship in that case modelled attribution as a sequential, teleological and, crucially, *interpretative* enquiry. As the Supreme Court presented this approach:

> The primary rule is contained in the company's constitution, [and/or] its articles of association ... But this will not cover the whole field of the company's decision-making. For this, the ordinary rules of agency and vicarious liability ... will normally supply the answer. However there will be some particular rules of law to which neither of these principles supplies the answer. The question is not then one of metaphysics but of construction of the particular rule in question.[37]

Daiwa had relied on *Bilta (UK) Ltd v Nazir (No 2)*[38] and, ultimately, *Stone & Rolls Ltd v Moore Stephens*[39] as authority for the proposition that the actions of a sole 'directing mind and will' (as they characterised Al Sanea) must be attributed to the company in the context of a claim such as this, even if his actions would not otherwise be so attributed. This obliged the Supreme Court to restate the reasoning of the seven-member panel in *Bilta (No 2)*, and clarify the status of *Stone & Rolls* in the wake of that decision.

In *Bilta (No 2)*, the Supreme Court 'held unanimously that where a company has been the victim of wrongdoing by its directors, the wrongdoing of the directors cannot be attributed to the company as a defence to a claim brought against the directors' in the name of the company. In contrast to this situation, Daiwa did not concern a claim against a fraudulent director, but against a third party who sought to rely on the illegality defence by attributing the director's default to the claimant company, recalling the facts of *Stone & Rolls*. With respect to such situations, Lords Sumption and Neuberger in *Bilta (No 2)* (with whom Lords Clarke and Carnwath concurred) understood *Stone & Rolls* to stand for two propositions. First, illegality can never be invoked by a third party (such as Daiwa) to defeat the company's claim against it if there are innocent shareholders or directors.

31 *Prest v Petrodel Resources Ltd* [2013] UKSC 34, [2013] 2 AC 415, [66] per Lord Sumption.

32 High authority had previously questioned (*obiter*) whether the veil could *ever* validly be disregarded: *VTB Capital plc v Nutritek International Corp* [2013] UKSC 5, [2013] 2 AC 337.

33 *Prest* (n 31) [8].

34 See, for instance, *Rossendale Borough Council v Hurstwood Properties Ltd* [2019] EWCA Civ 364, where separate personality was respected, despite its use for widespread evasion of rate payment.

35 *Daiwa* (n 1) [28].

36 [1995] 2 AC 500, 506–507.

37 *Daiwa* (n 1) [28].

38 [2015] UKSC 23, [2016] AC 1 (reported as *Jetivia SA v Bilta (UK) Ltd*).

39 [2009] UKHL 39, [2009] 1 AC 1391.

Second, the defence is *sometimes* available provided there are no such innocent shareholders or directors. The second proposition that a majority in *Bilta (No 2)* felt able to extract from *Stone & Rolls* appears to have propagated a belief in 'a rule of law that the dishonesty of the controlling mind in a "one-man company" could be attributed to the company ... whatever the context and purpose of the attribution in question'.[40] The Supreme Court in *Daiwa* emphatically denied any support in *Bilta (No 2)* for such a rule, endorsing instead Rose J's interpretation of the case: 'the answer to any question whether to attribute the knowledge of the fraudulent director to the company is always to be found in consideration of the context and the purpose for which the attribution is relevant'.[41]

The court in *Bilta (No 2)* – at least those members who thought *Stone & Rolls* stood for anything beyond its own facts – presented its contribution as a *negative* rule about attribution in the context of illegality, rather than a positive one. The presence of innocent actors precludes the defence, but their absence does not alone secure it. Beyond that negative rule, the question of attribution in the context of illegality is always a purposive, context-sensitive enquiry. The court in *Daiwa* therefore emphasised that, while the enquiry may resolve differently depending on whether its purpose is to apportion responsibility 'between the company and its agents' or 'between the company and a third party',[42] there is no *automatic* answer in the latter situation, even in the case of a so-called one-man company.

Its many critics have been awaiting the purging of *Stone & Rolls* from the law of attribution and illegality for some time. Already in *Bilta (No 2)* it attracted damning judicial passive aggression, characterised as a decision that 'stands as authority … that on the facts of that case no claim lay against the auditors, but nothing more'.[43] For Lord Neuberger it was 'not in the interests of the future clarity of the law for it to be treated as authoritative or [generally] of assistance'.[44] Nonetheless the fact that the court felt able to deduce at least some enduring principle from *Stone & Rolls* – the negative rule identified above – may have stayed its execution. The court in *Daiwa* was less merciful and has brought welcome clarity to the law by announcing that '*Stone & Rolls* can finally be laid to rest.'[45]

5 The relationship between attribution and illegality

Concluding its judgment, the Supreme Court perceived a simple case at the heart of the prolonged dispute:

> When it appeared that the company was running into difficulties, its 'directing mind' and sole shareholder fraudulently deprived the company of ... money by directing Daiwa to pay it away. Daiwa should have realised that something suspicious was going on and suspended payment until it had made reasonable enquiries to satisfy itself that the payments were properly to be made. The company ... has been the victim of Daiwa's negligence.[46]

40 *Daiwa* (n 1) [33].

41 *Daiwa* (n 4) [182] (endorsed by the Supreme Court: n 1 above [34]).

42 *Daiwa* (n 1) [30].

43 *Bilta (No 2)* n 38 [154], per Lords Toulson and Hodge.

44 Ibid [30], per Lord Neuberger. See further Ernest Lim, 'Attribution and the illegality defence' (2016) 79(3) Modern Law Review 476.

45 *Daiwa* (n 1) [34].

46 Ibid [39].

There is an explanation for why a case the court rightly accepted *post mortem* as 'bristling with simplicity'[47] nevertheless produced lengthy and complex judgments below and ultimately required resolution in the Supreme Court. That explanation lies in the enduring confusion over the attribution question in the context of illegality.

The law on illegality has been in rapid flux over recent years due to a quick-fire succession of Supreme Court decisions promulgating frankly inconsistent accounts of the doctrine.[48] Even when an ostensible settlement finally emerged (in *Patel*), it was in the form of an open-textured, contextually specific and policy-laden approach which – judicial denials notwithstanding[49] – is hard to distinguish in practice from a discretion.[50] *Patel* left enduring questions about the interaction between its own policy-laden illegality test and other relevant legal doctrines, and the status of a large amount of prior case law.

The approach to illegality confirmed in *Patel* represented a profound shift in the law's understanding of the relationship between illegal action and private law claims. *Patel*'s test derived from Lord Wilson's opinion in *Hounga v Allen*.[51] Although some commentators had presented this as a narrow exception to the (then) authoritative approach in *Tinsley v Milligan*,[52] confined to cases of a similar character,[53] nothing in Lord Wilson's *ratio* required any such limitation, and subsequent decisions confirmed the generality of the new approach to illegality. Particularly, they confirmed its ability to collapse distinctions between causes of action in private law. In *Bilta (No 2)*, the Supreme Court held unanimously that the crimes of a company director also constituting a breach of fiduciary duty could not be attributed to the company when the liquidators sued the director for that breach. Their reasoning, however, differed markedly. For Lord Sumption, the rules of corporate attribution applied 'regardless of the nature of the claim or the parties involved'; in this type of case, an exception *integral* to those same rules denied attribution because it would defeat the object of the directors' duties.[54] The majority, however, held that '[t]he primary question for the court is whether [the company's] claim against the directors ... is barred by the doctrine of illegality', which meant there was 'no need … to get into the subject of attribution'.[55] The success or failure of the action turned on 'whether it is contrary to public policy' that the claim should succeed.[56] Ultimately, they held that there was a significant public interest in enforcing the duty owed by an insolvent

47 Ibid.

48 See James C Fisher, 'The *ex turpi causa* principle in *Hounga* and *Servier*' (2015) 78(5) Modern Law Review 854. Cf eg Nicholas Strauss, 'Ex turpi causa oritur actio?' (2016) 132 Law Quarterly Review 236 and Lim (n 44) 484. In response to the latter, see James C Fisher, 'The latest word on illegality' [2016] Lloyd's Maritime and Commercial Law Quarterly 483, 484–485, note 12.

49 See *Patel* (n 2) [175], per Lord Neuberger.

50 Lord Neuberger's insistence in *Patel* that the illegality rule does not create a discretion is curious in light of his earlier observation that 'once a judge is required to take into account a significant number of relevant factors, and the question of how much weight to give each of them is a matter for the judge, the difference between judgment and discretion is … pretty slight': ibid [173]. It remains likely that in practice judges will respond with 'a value judgment, by reference to a widely spread mélange of ingredients, about the overall "merits" or strengths, in a highly non-legal sense, of the respective claims of the public interest and of each of the parties': ibid [206], per Lord Mance. See further Fisher, 'The latest word on illegality' (n 48) 486–487.

51 *Hounga* (n 3).

52 [1994] 1 AC 340

53 See eg Alan Bogg and Sarah Green, 'Rights are not just for the virtuous: what Hounga means for the illegality defence in the discrimination torts' (2015) 44 Industrial Law Journal 101, 122.

54 *Bilta (No 2)* (n 38) [86].

55 Ibid [131].

56 Ibid [166].

company's directors to the creditors and, on that basis, declined to allow the illegality doctrine to obstruct the claim. On this approach, whenever illegality exists in the factual matrix, the attribution question is subsumed into the illegality test – namely whether public policy, holistically considered, demands the success of the claim, or its defeat. The fact that the action concerns breach of a director's duties becomes just one more ingredient – albeit an important one – in the cocktail of policy considerations that go into answering that most sweeping of questions.[57] As will be explained below, *Patel* similarly regarded the presence of illegality as an invitation to adjudicate directly by first-order considerations of public policy, rather than through 'dry' doctrinal formulae.

Daiwa is significant chiefly because it at least partially arrests this trajectory. On the one hand, the court accepted that attribution is always a purposive, contextual enquiry which 'has to be seen in the context of the possible defences to which it might give rise'.[58] In emphasising this element of *Meridian Global*, it differed markedly from Lord Sumption's more general understanding of attribution in *Bilta (No 2)*.[59] Because the 'starting point' is that directors' acts are not in law those of the company, the law always demands a specific, purposive justification for attribution, which cannot be generally assumed. But on the other hand, with respect to the *place* of the illegality rule in attribution cases, the court in *Daiwa* echoed Lord Sumption's minority position in *Bilta (No 2)*. The court approached the attribution and illegality enquiries as logically separable, cumulative questions, rather than deciding attribution *through* the policy balancing test. This is conspicuously distinct from the majority approach in *Bilta (No 2)* and, for that matter, with *Patel*.

In *Bilta (No 2)*, the majority was determined to integrate policy balancing into the decision even though the same outcome could be reached through narrower, technical reasoning. *Patel* confirmed that approach. Lords Sumption, Mance and Clarke allowed the claimant's action for restitution for total failure of basis on orthodox, *Tinsley* reliance grounds. The claimant did not have to invoke the illegal purpose of the transfer to establish a cause of action, which must therefore succeed regardless of the circumstantial criminality. But the majority thought it unprincipled to decide the case without integrating into its decision the criminal character of the parties' project. For the majority, the criminal character of the parties' arrangement must determine the success or failure of the claim, even though reference to illegality was unnecessary to establish a complete cause of action.[60]

The crucial point of *Patel* therefore lies in the conscious rejection of a long-standing normative assumption, namely that the law should try to distinguish between (i) illegality *constitutive* of a claim and (ii) illegality that is merely part of its history. *Tinsley* and its legacy regarded the latter as irrelevant to the parties' formal relationship in private law, even though it may in some moralistic sense taint their interaction.[61] *Patel* insists instead that

---

57 Lord Sumption rejected that approach because the illegality defence was a rule of law whose engagement was not informed by competing public policy considerations in particular cases: ibid [99]–[102].

58 *Daiwa* (n 1) [12].

59 Namely that attribution should be *prima facie* assumed and only subsequently scrutinised in light of pertinent exceptions. Lord Sumption had invoked *Meridian Global* in his *Bilta (No 2)* judgment, but his own approach was analytically distinct, subverting conventional company law analysis: see Lim (n 44) 479.

60 *Patel* (n 2) [139], per Lord Kerr: 'In this case, the formation of the contract, its purpose and its performance all involved illegality. Under the single reliance master rule, it is said that all of this can be ignored because it is not necessary [for the claimant] to rely on the terms of the agreement ... This cannot be the correct way in which to deal with the impact of illegality ... It is surely better and more principled to examine why illegality should or should not operate to deny [the claimant] a remedy.'

61 In discussion of *Bilta (No 2)*, Lim conversely suggests that it was in fact 'unnecessary to depart from *Tinsley* in order to support the flexible approach in *Hounga*' (n 44) 485.

illegality is *always* definitional of the parties' relationship in private law. Where illegality can be found, it is wrong – indeed, 'unprincipled' – to decide the dispute other than by reference to the public policy considerations at stake. Full fidelity to the logical consequences of *Patel* would therefore have seen the *Daiwa* court deploy a holistic, policy-attentive balancing exercise *even though* – as the court's actual *ratio* proved – the same outcome could be reached through orthodox rules of attribution. It is striking therefore that the court effectively sidelined the *Patel* illegality doctrine.

6 *Daiwa*'s place in the *Patel* settlement

Daiwa can be read as a signal that *Patel* will not be permitted entirely to collapse the structures and categories of private law into policy-based impressionism. Notably, none of the justices in *Daiwa* had sat in either *Hounga* or *Bilta (No 2)*, and only one (Lady Hale) had sat in *Patel*. It is tempting to infer that the balance of opinion on this most divisive of private law doctrines has once again shifted at the highest level, producing a Supreme Court suspicious of *Patel*'s anti-formalist potential. Certainly, *Daiwa*'s reasoning arrests the momentum that policy-based reasoning had gained across that trilogy of cases. Despite generally approving the approach of the Court of Appeal below, Lady Hale expressed 'reservations' about one of its suggestions in particular, namely that, in evaluating an illegality defence, 'an appellate court should only interfere if the first instance judge has proceeded on an erroneous legal basis, taken into account matters that were legally irrelevant, or failed to take into account matters that were legally relevant'.[62] Here, disguised as a technical question about the intensity of appellate supervision, lies the fundamental ideological division that continues to haunt and complicate the illegality principle. The Court of Appeal's language is that of review over a discretion. If the illegality principle operates as a (structured) discretion which does not admit of a single correct answer, then an essentially *procedural* supervision of first-instance decision-making indeed appears appropriate. But if applying the illegality rule is truly an exercise in formal legal analysis, 'an appellate court is as well placed to evaluate the arguments as is the trial judge',[63] and the judge's conclusions should be subject to appeal in the ordinary way. Lady Hale reiterated Lord Neuberger's declaration in *Patel* that assessing the effect of the illegality principle is 'not akin to the exercise of discretion'.[64] She tantalisingly referenced 'cases concerning the illegality defence pending in the Supreme Court', in which 'it should not be assumed that this court will endorse the approach of the Court of Appeal' on this question.[65]

All these pending cases concern *Patel*'s impact on the force of prior authority. In each, the Court of Appeal has reasoned conservatively. It has either rejected invitations to depart from prior precedent on the strength of *Patel*, or relegated *Patel* to a superficial role as the formal clothing for conclusions which in substance derive from cases which reflect a far narrower conception of the illegality principle.

In *Stoffel v Grondona*,[66] the defendant conveyancing solicitors had negligently failed to register the transfer of title to property the claimant had purchased using a fraudulently obtained mortgage loan. When the claimant defaulted on repayments and the mortgagee sought a money judgment against her, she sued the defendant solicitors for losses incurred due to the property's non-availability as security. The first instance trial pre-dated

62 *Daiwa* (n 8) [65].

63 *Daiwa* (n 1) [21].

64 *Patel* (n 2) above [175].

65 *Daiwa* (n 1) [21].

66 [2018] EWCA Civ 2031.

Patel, and the judge therefore applied the *Tinsley* reliance rule, upholding the claimant's action because she did not need to rely on the illegality of her fraudulent mortgage application to show a fully constituted action in negligence against the solicitors. The Court of Appeal applied instead the formula established in the (recently decided) *Patel*, but concluded that the public policy balancing test favoured allowing the claim. There was a strong public interest in enforcing solicitors' duties of care to their clients, and denying the claim would not *per se* advance the (likewise important) policy of preventing mortgage fraud. Barring the action would be a disproportionate response to the claimant's wrongdoing because (i) she herself had not profited from the illegality, and (ii) the mortgagee institution did not complain of the fraud. The final point is revealing. Concealed within the court's application of the *Patel* formula was the normative position so definitional of *Tinsley* and so unpersuasive to the majorities in *Bilta (No 2)* and *Patel* itself: claims should succeed if the illegality is merely a part of the history of the dispute, incidental to the parties' relationship as determined by specific, formal rules of private law. The potentially radical *Patel* structure may provide the form, but the substance of the decision derived from a *Tinsley*an, rather than *Patel*ian, vision of the relationship between public policy and private rights.

To similar effect is the Court of Appeal's decision in *Henderson v Dorset Healthcare University NHS Foundation Trust*.[67] The claimant sued the defendant Trust in negligence, seeking damages for various losses – the consequences of her committing manslaughter attributable to the negligent provision of psychiatric care. The Trust accepted liability but submitted that recovery was barred by the illegality rule. The Court of Appeal ruled, on the strength of *Clunis v Camden & Islington Heath Authority* and *Gray v Thames Trains*,[68] that the claimant's limited degree of personal culpability for the crime did not affect the application of the illegality rule. Those decisions had each barred the kinds of loss Ms Henderson was claiming irrespective of the claimants' precise degree of personal wrongdoing, binding the court to dismiss her claims unless there had been a subsequent Supreme Court decision with which those previous cases were inconsistent. The Court of Appeal understood *Patel* to have cemented Lord Wilson's 'flexible and nuanced' formula in *Hounga* – which required a court to ask 'what aspect of public policy founds the [illegality] defence and whether there is another aspect of public policy to which application of the defence would run counter' – as authoritative for claims in contract and unjust enrichment (the context of *Patel* itself). But it failed to 'discern in the majority judgments in *Patel* any suggestion that *Clunis* or *Gray* were wrongly decided or ... cannot stand with the reasoning in *Patel*'.[69]

Patel's strongest claim to have changed the trajectory of decisions in the Court of Appeal is in *XX v Whittington Hospital NHS Trust*.[70] The claimant had been left infertile by clinical negligence and sought to enter a commercial surrogacy arrangement in California, where such arrangements are lawful and binding, using her own and/or donated eggs. The trial judge had barred her action for the costs of such an arrangement on the strength of prior authority,[71] which bound him to regard crucial elements of the claimant's plans as contrary to public policy. The Court of Appeal conversely permitted recovery of the costs necessary to undertake a commercial surrogacy in either jurisdiction, using either

67 [2018] EWCA Civ 1841.

68 [1998] QB 978 and [2009] UKHL 33, [2009] AC 1339, respectively.

69 *Henderson* (n 67) [125].

70 [2018] EWCA Civ 2832.

71 *Briody v St Helen's and Knowsley Area Health Authority* [2002] QB 856.

her own or donated eggs,[72] overturning as unprincipled and outdated any distinction between those scenarios.[73] *Patel* was instrumental in persuading the court that the correct outcome could not automatically be determined from prior analogous case law. The judgment emphasised the elastic nature of public policy considerations,[74] and drew explicitly on *Patel* as demanding a re-evaluation of the policy basis of prior decisions. Further, barring recovery would be disproportionately severe on the claimant since it would effectively deprive her of the prospect of a biological family. Even here, however, too much should not be made of *Patel*'s contribution to a decision at odds with prior case law. Reproductive health is an overtly policy-sensitive area. There have been major changes in social values and practice even in recent years, as the Court of Appeal emphasised, and the policy conclusions at the root of prior case law may by now have attracted judicial reconsideration even without *Patel*. More crucially, the same outcome would have resulted from proper application of the reliance rule. As the court noted, the claimant was not contemplating anything that would render her criminally liable.[75] The relevant primary legislation – the Surrogacy Arrangements Act 1985 – relates only to commercial surrogacy in the UK and does not affect the legality of a British citizen engaging in it overseas.[76] Even for domestic arrangements, it criminalises only the conduct of commercial surrogacy business, not those who pay for the help of a surrogate.[77] *XX* is not a case of *Patel* imposing public policy reasoning onto what would otherwise be a technical decision detached from policy considerations, or producing an outcome that could not have resulted but for *Patel*.

The Court of Appeal's general approach in these cases implies judicial sensitivity to the risk that *Patel*, if left unchecked, could adversely affect the clarity and coherence of private law. *Daiwa* suggests that, when these cases are decided on final appeal, the Supreme Court will emphasise the enduring authority of pre-*Patel* case law, rather than the flexibility of the illegality principle in light of *Patel*.

Critics of *Patel* may welcome such a retrenchment in the law of illegality. Nonetheless, as a matter of strict law it is hard to see how *Patel* can really leave unscathed the authority of leading cases indebted to a very different vision of the illegality doctrine and its place in the law. The incompatibility is perhaps most manifest in *Daiwa* itself, but it is also apparent in *Henderson*, in which the Court of Appeal adopted an unnaturally restrictive understanding of *Patel*. The *Patel* illegality rule is, at a basic level, inconsistent with the understanding of illegality which provides the logical grounding for decisions like *Clunis* and *Gray*, notwithstanding the fact that the justices in *Patel* considered those older decisions unproblematic. Despite dutifully incanting Lord Mansfield's primordial statement of principle that the illegality doctrine 'is not designed to achieve justice between the parties', the majority in *Patel* declared it important to avoid an 'incongruous result in legal and moral terms'[78] and to 'strive for the most desirable policy outcome'.[79] This requires a court to 'tak[e] into account a range of factors'[80] to establish 'whether the

72 *Henderson* (n 67) [84].
73 Ibid [94], [105].
74 Ibid [64].
75 Ibid [68].
76 Ibid [55], [76].
77 Ibid [56].
78 *Patel* (n 2) [127].
79 Ibid [91].
80 Ibid.

public interest ... should result in denial of the relief claimed'.[81] The court declined 'to lay down a prescriptive or definitive list' of the relevant considerations,[82] but Lord Toulson suggested these would 'include the seriousness of the [claimant's] conduct ... whether it was intentional and whether there was marked disparity in the parties' respective culpability'.[83] Structural blindness to the claimant's moral culpability – a refusal to go 'behind the conviction' – made sense when the illegality rule was supposed to operate independently of policy and episodic justice. The *Tinsley* reliance rule was, after all, a reaction against the moralistic 'public conscience' test of 1980s Court of Appeal jurisprudence.[84] In *Les Laboratoires Servier v Apotex*, Lord Sumption summarised it as precluding any judicial power to 'apply the illegality defence or not according to the relative importance which they attach to the policy underlying it by comparison with desirability of allowing an otherwise sound claim to succeed'.[85] But that is explicitly what *Patel* requires courts to do. Ultimately, there can be only one authoritative test for the operation of the illegality rule, which cannot be both (i) a reliance enquiry blind to the claimant's precise degree of moral responsibility – the animating principle in *Clunis* and *Gray* – and (ii) a public policy balancing test intimately focused on factors such as the moral quality of the claimant's actions. The Supreme Court in *Daiwa* was tolerant of sensitivity to the extent of the parties' respective wrongdoing. That plainly accords with *Patel*, but ineluctably puts logical pressure on prior cases which deny the relevance of comparative subjective wrongdoing. Despite this, on the whole *Daiwa* suggests that the Supreme Court also regards *Patel* as the high-water mark in the ascendency of the wide view of the illegality doctrine.

Such a recent and high-profile decision as *Patel* – the decision, moreover, of an expanded panel of the Supreme Court – is unlikely to be directly impugned at the highest level for many years, but its disruptive potential is likely to be forestalled when these several illegality cases receive the attention of the current justices. It is tempting to associate the dissipation of *Patel*'s momentum with a more general turn in private law, in which the courts have moved to compartmentalise public policy reasoning – not actually excluding it, but ensuring it operates within the architecture of formal law. Most conspicuously, in *Robinson v Chief Constable of West Yorkshire Police*,[86] Lord Reed's leading judgment declared 'mistaken' the belief – deriving from the opinion of Lord Bridge in *Caparo Industries v Dickman*[87] – that 'the court will only impose a duty of care where it considers it fair, just and reasonable to do so on the particular facts'.[88] Rather, in considering the appropriateness of duties of care in novel situations, a court should look principally to the established authorities so as to develop the law 'incrementally and by analogy'. It would be 'unnecessary and inappropriate' and indeed 'a recipe for inconsistency and uncertainty' to ask again whether such duties were fair, just and reasonable, since this will anyway be an important part of analysing whether a proposed novel duty is relevantly analogous to one accepted in prior decisions.[89]

81 Ibid [109].

82 Ibid [107].

83 Ibid.

84 Leading examples include *Saunders v Edwards* [1987] 1 WLR 1116 and *Euro-Diam v Bathurst Ltd* [1988] 2 WLR 517.

85 [2014] UKSC 55, [2015] 1 AC 430, [99].

86 [2018] UKSC 4.

87 [1990] 2 AC 605.

88 *Robinson* (n 86) [21].

89 Ibid [26].

Conclusion

Daiwa took place in the shadow of a sudden and dramatic retheorisation of the purpose, justification and operation of the illegality doctrine in English law. It was fundamentally a simple dispute and the Supreme Court dealt impressively with various attempts to exploit confusion in this area of the law, correctly dispensing with each argument for attribution.

More generally though, *Daiwa* and other cases in *Patel*'s wake cast the higher courts as engaged in an oyster-like process – clothing a potential irritant in a familiar substance to prevent damage to the system from within. The 'irritant' in this metaphor is the bold, flexible and expansionist illegality doctrine that emerged out of a line of Supreme Court authority begun in *Hounga* and culminating in *Patel* – a doctrine that privileges what one dissenter derided as 'a widely spread mélange of ingredients, about the overall "merits" or strengths, in a highly non-legal sense, of the respective claims of the public interest and of each of the parties'.[90]

English law still seems unable, or unwilling, to decide with conviction what it wants the illegality rule to be and to do, and consequently how it should relate to the fundamental organising rules for specific causes of action. As soon as there emerged an illegality principle sufficiently malleable and expansive to displace other means of resolving disputes – just as *Bilta (No 2)* recast a case about attribution into one about illegality – the law baulks and retreats into orthodoxy. The courts seem now at pains to emphasise the continued authority of cases decided on plainly un-*Patel*ian logic. This may simply be the common law's dialectical evolution at work: the emergence of synthesis out of thesis (*Tinsley*) and antithesis (*Patel*). Whether the process will produce something of pearlescent refinement remains to be seen – more will be clear once the Supreme Court gives judgment in *Henderson*, *Stoffel*, and *XX*. But if the conclusion on which the law ultimately settles is that the new, flexible and policy-attentive illegality formula – so long demanded by *Tinsley*'s eminent detractors[91] – leaves intact the old leading cases and does not actually require courts to reach different conclusions, what precisely was the point?

90 *Patel* (n 2) [206], per Lord Mance.

91 See, generally, eg Law Commission, *The Illegality Defence: A Consultative Report* (Law Com Consultation Paper No 189, 2009).

NILQ 71(3): 401–428

Duress and loss of control: fear and anger in excusatory defences

GLENYS WILLIAMS*

Aberystwyth University

'Without fear we are not human.'[1]

Abstract

This article examines the role of the anger and fear emotions in the loss of control and duress defences and argues that, although fear is now included as a trigger in loss of control, priority is still given to anger as a triggering event. Furthermore, in duress, although fear is the overriding mental state of the duressee, it wrongly forms no part of the rationale of the defence at all.

Following a brief examination of both emotions, the article – individually with respect to each defence – considers issues relating to the (in)sufficiency of the objective element contained in the defences, specifically because neither properly take fear into account as a characteristic which should be attributed to the reasonable person, and then, to a lesser extent, what impact theoretical principles, such as mechanistic and evaluative approaches, have on the role emotion plays in both defences (if any). It is clear that none of these, nor indeed the relatively new discipline of neuroscience, examined in the penultimate section of the article, can tell us about the effects of emotion on decision-making, reasoning, control and responsibility, nor can they provide an answer as to how emotions – fear especially – can be properly incorporated into both defences.

Numerous emotion-based alternative solutions are disseminated, and, although no preference is expressed here, it is recommended, firstly, that fear should be more effectively incorporated into the loss of control defence and, secondly, that duress should include fear as a characteristic attributed to the reasonable person.

Keywords: emotion; characteristics; fear; anger; loss of control; duress; choice/character theory; neuroscience; responsibility.

Introduction

Emotions play a vital part in all aspects of our daily lives, and indeed it is probably true to say that they have some influence – to a greater or lesser degree – on all the decisions we make about the way we live, be they trivial or life-changing.[2] It has been said that

* I am very grateful to the reviewers for their constructive comments.

1 J Finder, *Extraordinary Powers* (Orion 1993) 346.

2 See, for example, E Y Drogin and R Marin, 'Extreme emotional disturbance (EED), heat of passion, and provocation: a jurisprudent science perspective' (2008) 36 Journal of Psychiatry and Law 133, 139.

emotions 'facilitate decision making';[3] they help us to deal with problems;[4] they define our goals and values;[5] and are important to our welfare and comfort.[6] In particular, some primary emotions, such as anger and fear, which stem from our evolutionary survival instinct, play a functional[7] role in, for example, readying us to 'respond to challenges and opportunities … and providing us with information about what is important and how we are faring with respect to our goals'.[8]

The anger and fear emotions also play a part (although not always explicitly), in three criminal law defences, notably loss of control[9] (a partial defence to murder), duress (a full defence to all offences *except* murder) and self-defence (a full defence to murder). Traditionally, the anger emotion has most clearly been seen in the former provocation – now loss of control defence, albeit fear is now an additional component there. Fear is also a fundamental component in duress and self-defence.[10] As such, not only is it imperative that persons charged with criminal offences should be clear as to how these emotionally charged defences apply, it is also important for all manner of social, political and legal reasons that the way in which the emotional elements therein are interpreted should be both transparent and consistent. Yet, although the law and emotion debate has been evident for the last 20 years,[11] it will be seen that there are still no firm answers which provide the criminal law with a coherent understanding of the emotions which not only underlie these defences, but which also govern the way in which individuals behave.

To begin with then, this article will briefly clarify what emotions are, and what features pertain to the fear and anger emotions particularly, before going on to look at the loss of control and duress defences. While it is appreciated that there are clear differences between the two, which it could be argued would thus warrant a different approach, these two are selected for a number of reasons: Firstly, they are generally considered to be

3 J J Gross, 'The emerging field of emotion regulation: an integrative view' (1998) 2(3) Review of General Psychology 271, 272.

4 A H Fischer and A S R Manstead, 'Social functions of emotion' in M Lewis, J M Haviland-Jones and L F Barrett (eds), *Handbook of Emotions* (3rd edn, Guilford Press 2008) 456.

5 L C Charland, 'Is Mr Spock mentally competent? Competence to consent and emotion' (1998) Philosophy, Psychiatry and Psychology 67, 73–75.

6 D M Kahan and M C Nussbaum, 'Two conceptions of emotion in criminal law' (1996) 96(2) Columbia Law Review 269, 286.

7 i.e. that emotions have developed for a specific function; A Öhman, 'Fear and anxiety', and Fischer and Manstead, both in *Handbook of Emotions* (n 4) 710 and 456 respectively.

8 Gross (n 3) 'Conclusions'.

9 Although Herring rightly writes that, while the traditional focus was on anger, 'there is … nothing in the requirements of loss of control that requires the defendant to be acting out of anger'. J Herring, *Criminal Law: Text, Cases and Materials* (8th edn, Oxford University Press 2018) 230.

10 A Reilly, 'The heart of the matter: emotion in criminal defences' (1997–1998) 29 Ottawa Law Review 117, 141.

11 P G Nestor 'In defense of free will: neuroscience and criminal responsibility' (2019) 65 International Journal of Law and Psychiatry 1, 1. See, for example, S A Bandes (ed), *The Passions of Law* (New York University Press 1999); R Grossi, 'Understanding law and emotion' (2015) 7(1) Emotion Review 55; K Abrams and H Keren, 'Who's afraid of law and the emotions?' (2009–2010) 94 Minnesota Law Review 1997; H Petersen, 'The language of emotions in the language of law' in H Petersen (ed), *Love and Law in Europe* (Dartmouth 1998); C Sanger, 'Legislating with affect: emotion and legislative law making' in J E Fleming (ed), *Passions and Emotions* (New York University Press 2013).

excusatory, while self-defence is traditionally considered to be justificatory.[12] This – together with the absence of a need for proportionality in loss of control and duress – are the main reasons why self-defence will not be discussed further here. Also, as both loss of control and duress adopt an objective test which specifically imbues what was in the past described as the 'reasonable man' with a defendant's characteristics, a person's typical reaction to fear or anger is a characteristic which should thus, ostensibly at least, be taken into account. Thirdly, both loss of control and duress specifically look to 'circumstances' – especially that a fear of serious violence exists – to provide an explanation of the perpetrator's conduct. Next, and as will be discussed below, both duress and loss of control are 'concessions to human frailty'.[13] Sixth, they both involve an element of loss of control – a specific requirement in the loss of control partial defence, and in the case of duress, a loss of control by the defendant as a result of the pressure imposed by the duressor.[14] Finally, as the main aim of this article is to highlight the lack of regard given to the fear emotion, loss of control best demonstrates that, although fear of serious violence is now a triggering factor in loss of control, priority is still, as it always has been, given to the anger emotion,[15] whereas examining duress shows that, although fear is the overriding mental state of someone who is claiming duress,[16] it ironically forms no part of its rationale at all.[17]

Following the brief exposition of the fear and anger emotions, the article moves on to examine loss of control and, while acknowledging the role of characteristics and emotions in this partial defence, contends that anger – its traditional basis – takes precedence over the fear trigger, which is not given the priority it deserves.

The subsequent discussion of duress and its relevant characteristics shows that fear – its key component – is not taken onto account at all. In respect of both defences, this raises the question of the significance of the objective element and, to a lesser extent, the usefulness of both choice and/or character theory. What none of these do is explain whether the perpetrator was capable of controlling her behaviour or of resisting the threat, or whether she simply chose not to exercise that capacity.

---

12 See, for example, G Williams, 'The theory of excuses' [1982] Criminal Law Review 732, 734 and J Horder, 'On the Irrelevance of Motive in Criminal Law' in J Horder (ed), *Oxford Essays in Jurisprudence* (4th series, Oxford University Press 2000). However, this is not universally agreed. See, for example, Law Commission, *Murder, Manslaughter and Infanticide* (Law Com No 304, 2006) [6.61] with regard to duress; J Dressler, 'Rethinking heat of passion: a defense in search of a rationale' (1982) 73 (2) Journal of Criminal Law and Criminology 421, 448, with regard to provocation (as it then was) and R A Duff ,'Criminal responsibility and emotions: if fear and anger can exculpate, why not compassion?' (2015) 58(2) Journal of Philosophy 189.

13 See, for example, K J Arenson, 'The paradox of disallowing duress as a defence to murder' (2014) 78(1) Journal of Criminal Law 65, 71; and see, for example, the Irish Law Commission referred to in E Spain, *The Role of Emotions in Criminal Law Defences* (Cambridge University Press 2011) 186 and Lord Hailsham *R v Howe* [1987] 1 AC 417, 432. Or as the Law Commission has said: it is a 'recognition of the infirmity of human nature': *Codification of the Criminal Law. General Principles. Defences of General Application* (Law Com Working Paper No 55, 1974) [25].

14 The essence of duress is that '"pressure" or duress exerted upon a party in terms of threats [amounts to] submission …': J A Scutt, 'Consent versus submission: threats and the element of fear in rape' [1977] 13(1) University of Western Australia Law Review 52. This is borne out by a definition of submission as being 'a state in which people can no longer do what they want to do because they have been brought under the control of someone else': *Collins English Dictionary.*

15 See, for example, M J Allen, 'Provocation's reasonable man: a plea for self-control' (2000) 64(2) Journal of Criminal Law 216, 239; and A Clough 'Mercy killing: three's a crowd?' (2015) Journal of Criminal Law 358, 359.

16 Spain (n 13) 67, quoting Yeo.

17 E Spain, 'Duress and necessity in Ireland: reform on the horizon' (2008) 18 (3) Irish Criminal Law Journal 70, 71. See also *Simester and Sullivan's Criminal Law. Theory and Doctrine* (7th edn, Hart 2019) 753.

One way in which it could be possible to ascertain this is via the relatively new discipline of neuroscience, which forms the penultimate section of this article. While there will be a brief overview of some of the more recent emotion theories,[18] 'neurolaw',[19] the new 'rapidly developing area of interdisciplinary research on the meaning and implications of neuroscience for the law and legal practices',[20] will be analysed to ascertain if it is able to provide a means of explaining the impact these emotions have on cognitive decision-making and reasoning, and on a person's ability to exercise self-control (or not) or to demonstrate courage (or not), in the face of very strong emotions.

The conclusion reiterates that fear should be more effectively incorporated into the loss of control partial defence and should be a specific component of the duress defence. Some recommendations for emotion-based alternatives are highlighted in the conclusion.

1 Defining emotion[21]

The word emotion is based on the Latin *emovere* (*e* meaning 'out' and *movere* meaning 'move'). The related term 'motivation' is also derived from *movere*. In effect, there is a strong link between motive and emotion, as emotions are seen to 'motivate behaviour'.[22]

A dictionary definition describes emotion as 'a strong feeling deriving from one's circumstances, mood or relationship with others; instinctive or intuitive feeling as distinguished from reasoning or knowledge'.[23] However, there is in reality a general perception among emotion theorists that it is not possible to advance a coherent definition of emotion,[24] not least because it has such 'fuzzy boundaries',[25] and covers such a broad range of experiences. As Svendsen has explained: '"Emotion" is a term that can cover a range of highly dissimilar phenomena – from pain, hunger and thirst to pride, envy and love, from the almost purely physiological to the almost completely cognitive.'[26]

---

18 So not, for example, William James; Descartes: Freud. For more details on these, see D Keltner, K Oatley and J M Jenkins, *Understanding Emotions* (3rd edn, John Wiley & Sons 2014); and W Lyons, *Emotion* (Cambridge University Press 1980). A good summary of the various theories of emotion, and bibliography relating thereto can be found in the *Stanford Encyclopedia of Philosophy* <http://plato.stanford.edu/entries/emotion>.

19 See W Glannon, 'Neuroscience, law, and ethics' (2019) 65 International Journal of Law and Psychiatry 1 (Special Issue containing a number of articles on law and neuroscience); and O D Jones, R Marois et al, 'Law and neuroscience' (2013) 33 (45) 17 Journal of Neuroscience 624.

20 G Meynen, 'Neurolaw: neuroscience, ethics and law. review essay' (2014) 17 Ethical Theory and Moral Practice 819, 'Abstract'.

21 There are over 100 emotions listed by Curriculum Press <www.curriculumpress.edu.au/soi/downloads/TLI9_A-Z_Emotions.doc>.

22 Kahan and Nussbaum (n 6) 297. See the claimed re-emergence of the motive–emotion link in R M Ryan, 'Motivation and emotion: a new look and approach to two reemerging fields' (2007) 31 Motivation and Emotion 1.

23 *Oxford Dictionaries* <www.oxforddictionaries.com>. This is interesting for two reasons. Firstly, because emotion and instinct are two distinct concepts: see C Sanger, 'The role and reality of emotions in law' (2001–2002) 8 William and Mary Journal of Women and the Law 107, 108. Secondly, because of the assumption in the definition that emotion is separate from reasoning.

24 See, for example, J Hillman, *Emotion: A Comprehensive Phenomenology of Theories and their Meaning for Therapy* (Routledge & Kegan Paul 1960) 243; and R C Solomon, 'The philosophy of emotions' in *Handbook of Emotions* (n 4) 4.

25 J J Gross, 'Emotion Regulation;' in *Handbook of Emotions* (n 4) 498.

26 L Svendsen, *A Philosophy of Fear* (Reaktion Books 2008) 21.

Despite this ambivalence, there is partial concurrence that emotions can be divided into certain categories. Although there is no agreement as to an exact figure,[27] it has been claimed that there are six primary 'innate',[28] basic or 'simple'[29] emotions. These are 'happiness, sadness, fear, anger, surprise, [and] disgust'.[30] It is argued by some that we can have very little control over these because our reactions to them are both reflexive and instinctive.[31]

There are also purported secondary, 'complex',[32] 'developed'[33] or social emotions, such as embarrassment, love, shame, envy, guilt, pride, jealousy, awe and horror. These have also been described as 'higher cognitive emotions'[34] because they involve more reflection than the supposedly automatic primary emotions and, as such, are more liable to be swayed by our thoughts. Moreover, these take more time to develop and to recede than the primary emotions.[35] Of interest here is that a number of the secondary emotions seem to derive from the primary passions. For example, jealousy, awe and horror all include features of fear,[36] while guilt, pride, envy and shame all contain elements of disgust.

Finally, a third category of '*background*' emotions, comprises emotions such as 'well-being or malaise, calm or tension'.[37] Although similar to moods, the latter last for longer than emotions; they operate in a moderate, as opposed to intense ambience and are 'objectless, free- floating'.[38] Contrarily, emotions tend to endure for only a short time and always have an intentional object;[39] 'they are always about something or other. One is always angry about something … one is always afraid of something.'[40] If there was no object, there would be no reason to be either fearful[41] or angry.[42] As a result, there is available to us an underlying reason for responding to the emotion; there is a causal incentive to explain why the defendant acted in the way he did which is open to evaluation and as to its reasonableness or unreasonableness.[43]

27 D Evans, *Emotion: The Science of Sentiment* (Oxford University Press 2001) 6. Svendsen (n 26) notes that '[i]n an overview of fourteen lists of "basic emotions" it is striking that there is not one single emotion that is included in all lists' 22.

28 S Uniacke, 'Emotional excuses' (2007) 26(1) Law and Philosophy 95, 99.

29 J Bourke, *Fear. A Cultural History* (Virago Press 2005) 8.

30 A Damasio, *The Feeling of What Happens: Body, Emotion and the Making of Consciousness* (Vintage 2000) 50.

31 See, for example, J Le Doux, *The Emotional Brain. The Mysterious Underpinnings of Emotional Life* (Phoenix 1998) 19; and Evans (n 27) 16.

32 Bourke (n 29) 8.

33 Uniacke (n 28) 99.

34 Or even 'self-conscious emotions', A A Baird, 'The developmental neuroscience of criminal behavior' in N A Farahany (ed), *The Impact of Behavioral Sciences on Criminal Law* (Oxford University Press 2009) 95.

35 Evans (n 27) 28-9.

36 Bourke (n 29) 8.

37 Damasio (n 30) 51. Italics in original.

38 Keltner et al (n 18) 28.

39 As compared to moods; R S Lazarus, *Emotion and Adaptation* (Oxford University Press 1991) 48.

40 Solomon (n 24) 12.

41 Svendsen (n 26) 35.

42 S Gough, 'Taking the heat out of provocation' (1999) Oxford Journal of Legal Studies 19 (3) 481, 489.

43 J Gardner, 'The logic of excuses and the rationality of emotions' (2009) 43(3) Journal of Value Inquiry 1, 42.

This 'evaluative' view is one favoured by Kahan and Nussbaum[44] in their seminal article on two opposing approaches to emotions, the mechanistic and the evaluative.[45] The former refers to situations of impulse which are lacking in thought or cognition, deriving 'from an innate human nature' which impel a person to act impulsively.[46] Contrarily, the evaluative view is associated with a cognitive appraisal of events which are open to evaluation. This holds that emotions (i) possess within them an assessment or appraisal of the importance or significance of objects and events such that actors can themselves evaluate the appropriateness of their emotion and their resulting actions;[47] and also, by implication, (ii) that they 'can themselves *be* evaluated' by others for their appropriateness or inappropriateness[48] by judging cases of loss of control and duress alongside the way those emotions are typically expressed.[49] Therefore, while the evaluative view 'can appraise the evaluations internal to the offender's emotions as reasonable or not reasonable', the mechanistic view simply permits an inquiry into how resilient some emotions are.[50] In arguing that there is a cognitive element in virtually all emotions – including anger and fear – and that these are thus open to evaluation,[51] Kahan and Nussbaum have set the scene for the currently dominant evaluative/cognitive view.[52]

1.1 What is fear?[53]

Fear, 'an inevitable part of human existence',[54] is generally described as 'an emotional reaction resulting from the apprehension that there is a danger about and the consequent desire to avoid or be rid of the danger'.[55] It is a survival mechanism, designed to motivate behaviour to reduce any threats,[56] and, as one of the earliest existing emotions, has

44 Kahan and Nussbaum (n 6).

45 Legal theory is very sparse as most of the vast literature is mainly of a physiological, biological, psychological and psychoanalytic nature, as evidenced in the writings of, for example, Amélie Oksenberg Rorty, Joseph Le Doux and Ronald de Sousa.

46 Kahan and Nussbaum (n 6) 273 and 277–279. Mackay and Mitchell suggest that 'the reality lies somewhere in between' the evaluative and mechanistic view; R D Mackay and B J Mitchell 'But is this provocation? Some thoughts on the Law Commission's report on partial defences to murder' [2005] Criminal Law Review 44, 49. Having said that, research does show that even the more mechanistic/impulsive emotions do involve an element of subconscious processing. As Davies has said: 'despite our convictions that we consciously control our actions … [they] are in fact caused by lower level brain processes of which we have little or no awareness': P S Davies, 'Skepticism concerning human agency: sciences of the self versus voluntariness in the law' in N A Vincent (ed), *Neuroscience and Legal Responsibility* (Oxford University Press 2013). See also J Le Doux 'Coming to terms with fear' (2014) 111(8) Proceedings of the National Academy of Sciences USA 2871, and text to (n 197, n 198 and n 219) below.

47 i.e. that a 'subject's own evaluations of her situation were an essential part of her emotions': Kahan and Nussbaum (n 6) 286 and 291.

48 Ibid 286–287. Emphasis in original.

49 V Tadros, 'The Characters of Excuse' (2001) 21 (3) Oxford Journal of Legal Studies 495, 500.

50 Kahan and Nussbaum (n 6) 359–361. Compare Berger on what he calls 'normative veiling' i.e. simply saying that a person's will is overcome by emotion veils 'the normative quality of the emotions that inform the decisions that people make' and that they should, accordingly, be evaluated: B L Berger, 'Emotions and the veil of voluntarism: the loss of judgment in Canadian criminal defences' (2005-6) 51 McGill Law Journal 99, 128.

51 Kahan and Nussbaum (n 6) 295.

52 See, for example, Law Commission, *Report on Partial Defences to Murder* (Law Com No 290, 2004) [3.39].

53 According to Campbell, the '5 types of fear [are] (social; agoraphobic; death and illness; sexual and aggressive scenes; and harmless animals'): A Campbell, 'Sex differences in direct aggression: what are the psychological mediators?' (2006) 11(3) Aggression and Violent Behavior 237, 242.

54 Öhman (n 7) 709.

55 Lyons (n 18) 70.

56 Öhman (n 7) 710.

possibly been around for some 500 million years.[57] Consequently, we know more about this emotion than any other.[58]

Physically and physiologically while in a state of fear, muscles tense up, the eyes widen and the body perspires, yet is cool; both the heart rate and heartbeat accelerate. Some people may shake, run away, scream, jump, or simply curl up[59] and remain immobilised by the emotion. Indeed, usually, the automatic response to fear is to freeze.[60] In Galenic terms, this emotion is blue and cold.[61]

1.2 What is anger?

Anger has been defined as 'a strong feeling of annoyance, displeasure, or hostility';[62] 'a feeling of great annoyance or antagonism as the result of some real or supposed grievance'.[63]

However, this does not really convey the potential severity of anger reactions in loss of control situations especially; indeed, this is better conveyed in its synonyms, which suggest a greater level of anger, described in terms of rage, wrath, outrage and fury.[64]

When angry, one's violent propensities are heightened and concern about one's potential victim tends to diminish.[65] A person can be stronger, more aware and more forceful[66] when angry. 'The blood "boils", the face becomes hot, the muscles tense. There is a feeling of power and an impulse to strike out, to attack the source of anger … There is a strong feeling of impulsiveness and the "dimension of control" is lower than for any other emotion.'[67] In contrast to fear, the usual response to anger is to fight.[68] In Galenic terms, this emotion is red and hot. Contrasting the two emotions further, while it has been argued that 'anger … can be an ethically appropriate emotion and that … it may be a sign of moral weakness or human coldness *not* to feel anger',[69] the mainstream view is that anger is perceived to be an inappropriate or 'bad' emotion;[70] it is not usually perceived to be commendable, nor is it an emotion that we would want to encourage.[71] Illustrating this and writing about loss of control specifically, Simester and Sullivan have observed that 'anger is a criminogenic emotion and one of the key deterrent tasks of the

---

57 Evans (n 27) 44.

58 J Le Doux, *Synaptic Self* (Viking Penguin 2002) 212. Le Doux also claims that '[w]e have English words to distinguish more than three dozen variants of fear-related experiences'; Le Doux (n 46).

59 Hillman (n 24) 126; but see, for example, D H Barlow, B F Chorpita and J Turovsky, 'Fear, panic, anxiety and disorders of emotion' in D A Hope (ed), *Perspectives on Anxiety, Panic and Fear* vol 43 *Nebraska Symposium on Motivation* (Nebraska Press 1996) 289.

60 Le Doux (n 31) 176.

61 '(Medicine) of or relating to Galen (Latin name Claudius Galenus ?130-?200 AD), the Greek physician, anatomist, and physiologist, or his teachings or methods' <www.thefreedictionary.com/Galenic>.

62 Oxford Dictionaries <www.oxforddictionaries.com>.

63 Ibid.

64 Ibid. On anger and degrees of anger generally, see J Horder, *Provocation and Responsibility* (Clarendon Press 1992) and Dressler (n 12) 465.

65 R Brandt, 'A motivational theory of excuses in the criminal law' in M L Corrado (ed), *Justification and Excuse in the Criminal Law. A Collection of Essays* (Garland 1994) 114.

66 E A Posner, 'Law and the emotions' (Working Paper 103, University of Chicago) 5.

67 Reilly (n 10) 133, quoting Izard.

68 Gough (n 42) 487.

69 i.e. righteous anger: Law Commission (n 52) [3.38]. Emphasis added.

70 See, generally, B Rosebury, 'On punishing emotions' (2003) 16(1) Ratio Juris 37.

71 E Spain, 'Love in life and death' (2013) 64 (1) Northern Ireland Legal Quarterly 91, 106.

criminal law is to provide incentives for citizens to curb violent responses induced by their angry states'.[72]

On the other hand, the fear emotion has been said to be 'understandable … admirable',[73] 'appropriate and justified'.[74] It has been said that an emotion can be 'appropriate' where the reason for it is a good reason. As Charland illustrates, if a person came across a big bear in the forest, this would be a good reason for fear to be explained as both a cogent and fitting response.[75] Thus, acting to save oneself in such a situation would be 'a good reason to act which ought not to be ignored'.[76] In contrast, a 'petty affront' would not be sufficient.[77] This is reiterated in *R v Dawes*, where Lord Judge CJ said that 'unless the circumstances are extremely grave, normal irritation, or even serious anger, do not often cross the threshold into loss of control'.[78]

What we can see here then is that appropriateness is linked with the reasons for acting (taking us back to the link with motive) and that not only must the emotion itself be appropriate, but also that a person's response or reaction to it must also be appropriate. Uniacke explains that

> … an emotional response is of a morally appropriate type if a morally well-disposed person would be justified in having that type of response to the particular circumstances ... A morally appropriate type of response could be one that we would be right to feel in particular circumstances, such that it would be morally inappropriate were we to feel otherwise.[79]

Thus, while fear would be an appropriate response in situations of abuse especially, a response made in anger may *not* be appropriate. For example, Holton and Shute, writing about the previous provocation defence, said that '[i]t is one thing to get angry; it is another to lose one's self-control … Perhaps it is justifiable to become angry in the face of provocation. But is it justifiable to lose one's self-control? We think not.'[80] In the same vein, Norrie, agreeing with the Law Commission, wrote that 'anger cannot justify outright a violent response, certainly not a killing'.[81]

Despite these clear differences in our perception of the two emotions and our reactions to them, it should be noted that they do nonetheless have some common features. For example, the impulse, or instinct, of self-preservation is a central feature of

---

72 *Simester and Sullivan* (n 17) 411. Thus, raising the broader question of whether *any* defence based on anger or loss of control is acceptable; for example, Horder, referring to the then provocation defence, asked 'whether [any] actions in anger are worthy of excuse': J Horder, 'Assisting in suicide. keeping the debate alive' (1990) 54(1) Journal of Criminal Law 253 at 255. Also see Horder (n 64). Thanks go to a reviewer for raising this broader point.

73 J Horder, 'Cognition, emotion, and criminal culpability' (1990) Law Quarterly Review 469, 480. Uniacke has said that fear is a proper and fitting sort of emotional reaction to allow an excusatory defence: Uniacke (n 28) 101.

74 A Norrie, 'The Coroners and Justice Act 2009 – partial defences to murder (1) loss of control' [2010] Criminal Law Review 275, 278.

75 Charland (n 5) 77.

76 Reilly (n 10) 145.

77 *Per* Lord Millett in *R v Smith (Morgan)* [2001] 1 AC 146, 214.

78 *R v Dawes* [2013] EWCA Crim 322 [60].

79 'Lack of pity, for instance, would be callous or grossly insensitive in some circumstances'; Uniacke (n 28) 100.

80 R Holton and S Shute, 'Self-control in the modern provocation defence' (2007) 27 (1) Oxford Journal of Legal Studies 49, 70.

81 A Norrie, *Crime Reason and History* (3rd edn, Cambridge University Press 2014) 312 and Law Commission, *Partial Defences to Murder* (Law Com Consultation Paper No 173 2003) [4.165].

both emotions;[82] in an anger situation fighting, or other like hostile behaviour, can be seen as key to survival,[83] as acting in duress is, if the individual's life is threatened.[84] In addition, they are both primary (and negative) emotions, the responses to which may, to a degree, be reflexive, instinctive and reactive.[85]

2 The loss of control partial defence

In a series of papers in 2003,[86] 2004[87] and 2006,[88] the Law Commission reviewed the law on provocation and excessive use of force in self-defence (the latter in the context of providing a defence which did not discriminate against women who killed abusive partners, as provocation was deemed to do). In particular, and following criticism of the emphasis on anger as its focal point, it noted in its 2003 Consultation Paper, *Partial Defences to Murder*, that '[t]he defence of provocation elevates the emotion of sudden anger above the emotions of fear, despair, compassion and empathy, [and asked itself whether it was] morally sustainable for … anger to found a partial defence to murder'. In response, it concluded that there was an argument that it was not.[89] Furthermore, the Law Commission had also recommended that the need for a loss of control be dispensed with,[90] but this was rejected following consultation – and indeed, the nomenclature of loss of control replaced the provocation label.

With these issues in mind, the government subsequently acknowledged that ignoring the impact of fear in provocation was problematic and, as a result, the defence now contains a new 'fear of serious violence' 'trigger', although the Ministry of Justice itself conceded that 'it is not helpful for killings which are triggered primarily by fear to be shoehorned into a partial defence which is aimed at killings triggered by anger'.[91]

The provisions are contained in sections 54 and 55 of the Coroners and Justice Act 2009, which state that:

54

(1) Where a person ('D') kills or is a party to the killing of another ('V'), D is not to be convicted of murder if—

(a) D's acts and omissions in doing or being a party to the killing resulted from the D's loss of self-control,

(b) the loss of self-control had a qualifying trigger, and

82 Posner (n 66) 16.

83 Lyons (n 18) 41.

84 For example, the Law Commission (n 12) said: 'The law must recognize that the instinct and perhaps the duty of self-preservation is powerful and natural …' [6.64].

85 See, for example, R G Fontaine, 'The wrongfulness of wrongly interpreting wrongfulness: provocation interpretational bias and heat of passion homicide' (Selected Works of Reid G Fontaine, Bepress) 79 <http://works.bepress.com/reid_fontaine/11>.

86 Law Commission (n 81) (2003).

87 Law Commission (n 52), and indeed in Law Commission, *A New Homicide Act for England and Wales?* (Law Com Consultation Paper No 177, 2005).

88 Law Commission (n 12) (2006).

89 Law Commission (n 81) [4.164].

90 Law Commission (n 52) [1.13]. See, for example, B Mitchell, 'Loss of self-control Under the Criminal Justice Act 2009: oh no!' in A Reed and M Bohlander, (eds), *Loss of Control and Diminished Responsibility. Domestic, Comparative and International Perspectives* (Ashgate 2011) 50.

91 Ministry of Justice, *Murder, Manslaughter and Infanticide: Proposals for Reform of the Law* (MoJ CP19/08, 2008) [27].

(c) a person of D's sex and age, with a normal degree of tolerance and self-restraint and in the circumstances of D, might have reacted in the same or in a similar way to D …[92]

55 …

(2) A loss of control had a qualifying trigger if subsection (3), (4) or (5) applies.

(3) This subsection applies if D's loss of self-control was attributable to D's fear of serious violence from V against D or another identified person.

(4) This subsection applies if D's loss of self-control was attributable to a thing or things done or said (or both) which –

(a) constituted circumstances of an extremely grave character, and

(b) caused D to have a justifiable sense of being seriously wronged.

The Ministry of Justice explained that:

> Whichever of these qualifying triggers applies, a number of factors must be present for the defence to succeed.
>
> The first is that a person with certain characteristics might have acted in the same or similar way to the defendant. These characteristics are that: (i) they were of the same sex and age as the defendant; (ii) they had an ordinary level of tolerance and self-restraint; and (iii) they were in the same circumstances of the defendant …'[93]

2.1 Emotions and character(istics) in loss of control

As can be seen then, the relevant characteristics to be ascribed in loss of control are sex and age, tolerance and self-restraint,[94] and circumstances.[95] However, this is slightly misleading, because for the purposes of the objective test, it must be noted that, whereas a defendant's characteristics can be ascribed for the purposes of the gravity of the provocation, only age and gender – and no other characteristics – can be ascribed to the defendant's capacity for control.[96]

This distinction has been criticised most especially by Yeo on the basis that the rationale for the defence and the rationale for the objective test are two different things. According to Yeo, the rationale for the defence 'may be seen as the law's concession to human frailty… [or] that the accused was not fully in control of her or his behaviour when the homicide was committed'. However, the rationale for the objective test is 'the perceived need "for society to maintain objective standards of behaviour for the protection of human life"'. As such, the objective test rationale

> ... bears no conceivable relationship with the underlying rationale of the defence [which does not] require the distinction to be made between characteristics of

92 S 54(2) and (3) continue: '2) For the purposes of subsection (1)(a), it does not matter whether or not the loss of control was sudden.

(3) In subsection (1) (c) the reference to "the circumstances of D" is a reference to all of D's circumstances other than those whose only relevance to D's conduct is that they bear on D's general capacity for tolerance or self-restraint.'

93 Ministry of Justice Circular 2010/13, 'Partial defences to murder: loss of control and diminished responsibility; and infanticide. Implementation of Sections 52, 54 to 57 of the Coroners and Justice Act 2009' (4 October 2010) [16] and [17].

94 The allusion to self-restraint recognises that there is an expectation that people will exercise self-control. Herring (n 9) 240. Loss of control does not have to be absolute; see, for example, Lord Diplock in *Phillips v The Queen* [1969] 2 AC 130 cited in Law Commission (n 52) [3.27] on 'degrees' of control.

95 On circumstances especially, see C Withey, 'Loss of control, loss of opportunity?' [2011] Criminal Law Review 263, 273–274.

96 The same applied in the old provocation defence: *AG for Jersey v Holley* [2005] UKPC 23.

> the accused affecting the gravity of the provocation from those concerned with the power of self-control.[97]

It is a truism to say that the capacity for exercising self-control obviously differs as between different people, and while, for example, anger/fear management training[98] or emotion regulation[99] can be effective in training some individuals[100] to control their emotions,[101] some people will still be prone to reacting more aggressively than others.[102] This is no doubt one reason why characteristics cannot be ascribed to the person with a normal degree of tolerance and self-restraint in relation to the capacity for control, but does this mean that an inability to control oneself because one is angry is (or is not) a reflection of one's character?[103]

On the one hand, Reilly has said that an angry reaction *does* reveal a person's character because '[t]he person of good character has control over his or her emotional reactions',[104] while, on the other hand, character theory tells us that, even though the defendant was acting unlawfully, this was not a reflection of her character.[105] Tadros explains that this is because, in relying on reasons which are different to those that would normally motivate her, the actor is acting in a way which is distinct from her normal behaviour:

> Her character while she is in a state of extreme anger is not like her character whilst calm … But … this is so only if the settled character of the agent is

97 S Yeo, 'Power of self-control in provocation and automatism' (1992) 14 Sydney Law Review 3, 4 and 8.

98 These consist of 'behavioral, cognitive, attentional, physiological, or emotional strategies to eliminate, maintain, or change emotional experience and/or expression': L R Brody and J A Hall, 'Gender and emotion in context' in *Handbook of Emotions* (n 4) 400.

99 'Emotion regulation studies how individuals influence which emotions they have, when they have them, and how they experience and express them': Gross (n 3) 'Abstract'.

100 Such as, say, police officers, firefighters and soldiers. See, for example, M Baron 'Excuses, excuses'(2007) 1 Criminal Law and Philosophy 21, 23; See generally (2004) 17(1) Social Research: An International Quarterly of the Social Sciences (special 'Courage' issue); H V Hall, 'Extreme emotion' (1990) 12 University of Hawaii Law Review 39; and M D Bayles, 'Reconceptualising necessity and duress' in Corrado (n 65) 449. Grossman describes the 'conditioning' enforced on soldiers before going into battle, while Rachman uses bomb disposal experts (and astronauts) as examples of this; D Grossman, *On Killing. The Psychological Cost of Learning to Kill in War and Society* (Little, Brown & Co 1995); S J Rachman, 'Fear and courage: a psychological perspective' special 'Courage' issue 158–159. See also Morse's list of 'variables' that assist us in maintaining control: S J Morse, 'Culpability and control' (1993–1994) 142 University of Pennsylvania Law Review 1587, 1605–1610.

101 See, for example, Morse (n 100). In effect, this suggests that we are not talking about the emotion itself, but rather, a person's reaction to it. See text to (n 79) above. As Reilly (n 10) 133 has said, '[a]lthough actors might have no control over the arousal of emotion, they have a significant level of control over how to deal with the feelings invoked. The translation of anger into aggression is a matter of choice.' Compare J Sabini and M Silver, 'Emotions, responsibility and character' in F Schoeman, *Responsibility, Character and the Emotions* (Cambridge University Press 1987) 169 (summarising Kant's view on emotions). On controlling emotions, see, for example, Holton and Shute (n 80), and R Ingalese, *The History and Power of Mind*, chapter 4 'The art of self-control' (Dodd, Mead & Co 1929) <www.svpvril.com/Ingalese/ingalese4.html>.

102 B Mitchell, 'Provoked violence, capacity and criminal responsibility' (1995) 1 Psychology, Crime and Law 291, 296.

103 J T Parry, 'The virtue of necessity: reshaping culpability and the rule of law' (1999) 36 Houston Law Review 397, 423–426. G Williams, 'Necessity: duress of circumstances or moral involuntariness?' (2014) 43 1 Common Law World Review 1, 23.

104 Reilly (n 10) 132. N Levy and T Bayne, 'A will of one's own: consciousness, control and character' (2004) 27 International Journal of Law and Psychiatry 459, 468: 'it will remain true that some individuals will have less will power than others do through no fault of their own'.

105 Tadros notes this (n 49) 495.

> peaceful. If it is not, then she cannot show the difference between her settled character and her outraged character that provides the basis of the defence.[106]

This points towards some interesting philosophies about the effect of emotion on loss of control,[107] whereby it is argued that the provoked person is still capable of reasoning,[108] but her perspective is changed such that she prioritises reasons for action that would normally have remained marginal, rather than relying on reasons that would usually influence her. Posner best explains this as follows:

> Emotions … have a certain feel or affect characterized usually by a focus on particular stimulus with the result that the rest of the environment 'fades' (a little or a lot, depending on the strength of the emotion) … [During the emotion state] people continue to act rationally [albeit] differently from the way they do in the calm state … people experience … temporary variations in their preferences, abilities, and/or beliefs.
>
> Their preferences change so that what psychologists call the 'action tendency' of an emotion becomes relatively attractive. The action tendency of anger is to strike out; so we can say that a person, while angry, develops a temporary preference to strike the person who offends him … a person in an emotional state does not act irrationally given his temporary preferences … [in contrast] fear produces flight from a threat.[109]

It can be seen then that there is a disparity between the anger and fear reactions, and that the latter is not a reaction customarily associated with losing self-control.[110] In line with the typical responses to anger and fear noted earlier, the former characteristically manifests itself as a violent expression of rage, whereas the latter 'might be characterised by very different external signs which are not easily detectable: typically a state of paralysis and submission'[111] or, as noted by Posner, a flight reaction, none of which relate to the notion of loss of control.

This therefore speaks directly to the question as to whether it was necessary to keep the loss of control requirement in the defence. When the Law Commission recommended dispensing with it, Mackay and Mitchell argued that the loss of control prerequisite went to the very heart of the defence because it had to be caused by 'extreme emotional disturbance' on the defendant's part. Abolishing loss of control would thus ignore the defendant's mental condition in a situation where it was precisely that which caused him to lose his self-control in the first place.[112]

Contrarily, however, and agreeing with the Law Commission, Carline noted that:

> [t]he essence of the defence is fear and thus the law should not also require the defendant to suffer a loss of self control … If the Government consider that it is important to abolish the word 'provocation' because of its negative connotations it is difficult to understand why the same does not apply to the

106 Tadros ibid 507 and, more generally, in *Criminal Responsibility* (Oxford University Press 2007), Chapter 11.

107 By, for example, Gough (n 42) 481 and 488, and G Mousourakis, 'Emotion, choice and the rationale of the provocation defence' (1999) 30 Cambrian Law Review 21, 23.

108 In *R v Jewell* the Court of Appeal accepted that loss of control meant 'a loss of *normal* powers of reasoning' [2014] EWCA Crim 414 [23], emphasis added.

109 Posner (n 66) 3–5. Compare Holton and Shute (n 80) 52 and 55: 'What is lost when one loses self-control is control over *which* mental elements [and "immediate inclinations"] drive one's actions'.

110 Mitchell (n 90) 47.

111 Reilly (n 10) 137.

112 Mackay and Mitchell (n 46) 47.

> phrase loss of self control, which connotes anger as opposed to fear and desperation.[113]

This must be right. If we are saying that loss of control is an emotion-based defence and that allowances are given for reactions carried out as a result of emotions felt at the time the offence was committed, then why the need for an additional requirement in the form of loss of control which does not 'fit' with the fear emotion? The government's explanation that it needed to be retained in order to prevent its use in cold-blooded killings is fair enough, but this does prioritise that one form of killing at the expense of others, such as those committed as a result of domestic violence, for example, which could legitimately be claimed on an emotional rationale,[114] a rationale the government clearly wanted to avoid.

However, that criticism aside, and despite its welcome inclusion in the partial defence, it is also clear for other reasons that the fear trigger does not have the force of anger in loss of control.

Firstly, the conditions for satisfying the triggers are unequal. This is made evident by Susan Edwards who explains that:

> Fear … is only a necessary condition, it is not a sufficient condition and is qualified not by 'extremely grave circumstances', as is the anger defence, but specifically by the fear of 'serious violence' … so … the evidential requirement is more stringent and more specific than the anger qualifying trigger … The state of anger must follow on from extremely grave circumstances whilst the state of fear must flow from serious violence. In addition, extremely grave circumstances can include any circumstances, whilst there is one precondition to the fear defence, that of 'serious violence'.[115]

Secondly, the fear–loss of control trigger incorporates excessive use of force in self-defence. The Law Commission decided (and the government agreed) that, rather than have a separate defence of excessive force, this should be integrated into the loss of control partial defence, which would accordingly cover cases of *overreaction* to a fear of serious violence, where (i) excessive force is used, and (ii) the attack was not imminent enough to amount to self-defence.[116] One of the implications of this is that whereas fear is seen as an overreaction, anger is not. This is because the fear trigger accommodates cases where the defendant could not use self-defence because the amount of force used was unreasonably excessive. So, while on the one hand the fear trigger makes it easier for an abused person to take advantage of the defence, on the other, it fails to appreciate that hers is not an overreaction, and neither is it excessive.[117] On the contrary, and as Reilly has argued, the fear she felt was, for her, an endorsement of the reasonableness of her self-defensive conduct.[118]

---

113 A Carline, 'Reforming provocation; perspectives from the Law Commission and the government' [2009] 2 Web Journal of Current Legal Issues <http://webjcli.ncl.ac.uk/2009/issue2/carline2.html>.

114 More will be said in this point when looking at the MPC in the 'Conclusion'.

115 S S M Edwards, 'Loss of self-control: when his anger is worth more than her fear' in Reed and Bohlander (n 90) at 91.

116 Law Commission (n 52) [4.17], and Law Commission (n 12) [5.53–5]. Emphasis added.

117 The point is made by S Edwards, 'Anger and fear as justifiable preludes for loss of self-control' (2010) 74 (3) Journal of Criminal Law 223, 234.

118 Reilly (n 10) at 141. For further reading on the 'reasonableness of killing' see, for example, W Wilson, 'The structure of criminal defences' [2005] Criminal Law Review 108, 117: 'Killing, even in justified anger, is the antithesis of reasonableness' and Law Commission (n 81) [4.162]: provocation 'raises the question whether a reasonable person should ever respond to provocation by killing'.

Thirdly, fear of serious violence, and how it affects a person's behaviour within the context of abuse in particular, is a situation with which the average juror will not be familiar.[119] This makes it even more difficult to fit fear into the loss of control requirement and to compare it with the reaction of a person with a normal degree of tolerance and self-restraint.

3 The duress defence[120]

It seems that, in England and Wales, there are now two forms of duress – duress by threats and duress of circumstances. They are both common law defences, but not to murder. This has been the subject of much debate, and much to-ing and fro-ing on the part of the Law Commission,[121] which in 2006 eventually reverted to its original recommendation that duress should be a full defence to murder (and attempted murder)[122] on a rationale previously expressed that no 'social purpose is served by requiring the law to prescribe … standards of determination and heroism'[123] that were unattainable. The government did not take the Law Commission's duress proposals forward in its 2008 Consultation Paper on *Murder, Manslaughter and Infanticide: Proposals for Reform of the Law*; it simply did not look at duress, focusing only on provocation, diminished responsibility, complicity and infanticide.[124]

To satisfy duress by threats, the defendant must demonstrate that he committed the crime as a result of threats of death or grievous bodily harm and that a reasonable person sharing the characteristics of the defendant would have acted as he did.[125] Duress of circumstances provides a defence where the defendant reasonably believes that the *circumstances* are such that unless he or she commits a crime he or she or another will suffer death or serious injury and that 'a sober person of reasonable firmness, sharing the characteristics of the accused [would] have responded ... as [he did]'.[126] Thus it can be seen that '[t]he first limb is concerned with the nature of D's fear and the seriousness of the threat [while] [t]he second involves an evaluation of the reasonableness of response'.[127]

119 L Claydon and C Rödiger, 'Fear, loss of control and cognitive neuroscience' (2016) 22(2) European Journal of Current Legal Issues, or, as Loveless writes, one of the fundamental problems is 'general ignorance of the psychological effects of domestic violence': J Loveless, 'Domestic violence, coercion and duress' [2010] Criminal Law Review (2) 93, 100.

120 For a more detailed exposition of this duress section and of some of the points made below, see Williams (n 103) especially 9, 11, 19–20; and 22.

121 Law Commission, *Codification of the Criminal Law: General Principles – Defences of General Application* (Law Com Working Paper No 55, 1974) [25] recommended that it should be defence to murder; likewise in its 1977 *Criminal Law: Report on Defences of General Application* (Law Com No 83) [2.42]; *Criminal Law: A Criminal Code for England and Wales* (Law Com No 177, 1989) [12.13] vol 2 recommended that it *not* be a defence to murder (because of *Howe*). In *Legislating the Criminal Code: Offences Against the Person and General Principles* (Law Com Consultation Paper No 122, 1992) [18.14], it reverted back to a full defence. It maintained this view in *Legislating the Criminal Code. Offences Against the Person and General Principles* (Law Com No 218, Cmnd 2370, 1993) [29.14]. By 2005, in Law Commission (n 87) [1.42], it was recommending that duress should only reduce first to second degree murder.

122 Law Commission (n 12) [6.21].

123 Law Commission, *Report on Defences of General Application* (Law Com No 83, 1977) [2.43].

124 Ministry of Justice (n 91) [7] and [8].

125 Herring (n 9) 643. See *R v Graham* [1982] 1 WLR 294 confirmed in *R v Hasan* [2005] UKHL 22.

126 *R v Martin* [1989] 1 All ER 652 *per* Simon Brown J 654.

127 Loveless (n 119) 96.

3.1 Emotions and character(istics) in duress

In duress, the relevant characteristic to be taken into account is courage[128]/firmness[129] in the face of threats 'to life and limb'[130] because the overriding mental state of the individual being threatened is fear.[131] As a person who possesses courage is one who is able to overcome fear,[132] the link between fear and courage is thus clearly made,[133] yet this is not attributed to the reasonable man in the objective test at all. Indeed, Duff has contended that the reasonable man should be credited with 'any of this defendant's actual characteristics that affected his response to the threat, *other than characteristics which involve or reveal a lack of … proper courage*'.[134]

There are three problems with this. Firstly, there is an unrealistic assumption that a minimum standard of courage is expected of all individuals.[135] For example, in *R v Horne*[136] and in *R v Bowen*[137] it was held that any emotional characteristics such as vulnerability or susceptibility to threats suffered by the defendant should not be ascribed to the reasonable person nor be taken into account in the objective test. As Virgo has commented, the fact that Bowen's 'low intelligence' and that he was 'more pliable, vulnerable, timid or susceptible to the threats … did not make him less courageous than the reasonable person'.[138] The same could be said of coerced victims of domestic violence highlighted by Loveless. In an argument similar to that propounded by Susan Edwards on loss of control and domestic violence noted earlier, Loveless argues that because fear will always be present, it should be acknowledged in the subjective element of the test 'that victims of violence have a greater sensitivity to the risks in their environment than would be obvious to an observer' and that the objective element

> does not accommodate one whose ability to resist a threat is lower than that of a reasonable person … The relevance of this is that an abused woman who has [chosen to remain] with her abusive partner … is likely to be regarded as weak, submissive and vulnerable in the absence of any relevant characteristic.[139]

---

128 See, for example, G Virgo 'Are the defences of provocation, duress and self-defence consistent?' (2002) 4 Arch News 4: 'the function of this objective test is to impose a test of courage'.

129 K J M Smith, 'Duress and steadfastness: in pursuit of the unintelligible' (1999) 363, 372.

130 D Pears, 'The Anatomy of Courage' special 'Courage' issue (n 100) 6 and 7.

131 Spain (n 13) 67, quoting Yeo.

132 G Kateb, 'Courage as a virtue' in special 'Courage' issue (n 100) 43. Courage is defined as the 'mental or moral strength to … withstand danger, fear, or difficulty': Merriam-Webster <www.merriam-webster.com>.

133 Pears (n 130) 7.

134 A Duff, 'Choice, character and criminal liability' (1993) 12(4) Law and Philosophy 345, 359. Emphasis added.

135 See, for example, J Gardner, 'The gist of excuses' (1997–1998) 1 Buffalo Criminal Law Review 575; Berger (n 50); J Dressler ,'Exegesis of the law of duress: justifying the excuse and searching for its proper limits' in Corrado (n 65) 381.

136 *R v Horne* [1994] Crim LR 584, 585.

137 *R v Bowen* [1996] 4 All ER 837, 844.

138 Virgo (n 128).

139 Loveless (n 119) 98–99. Loveless argues elsewhere that duress should be extended to include coercion in the context of domestic violence; J Loveless, 'R v GAC: Battered Woman "Syndromization"' [2014] 9 Criminal Law Review 655, 662.

Secondly, in any event, the standard expected is not simply that of the reasonable man.[140] Rather, it is one of heroism.[141] This can be seen in Lord Coleridge and Lord Hailsham's *dicta* in *R v Dudley and Stephens*,[142] and in *R v Howe*[143] respectively, to the effect that a reasonable person must heroically sacrifice his life for others. Reilly has criticised Hailsham's judgment because, when he decided the defendants were cowards, he took them out of their social context and, instead, imposed his own point of view which was that of a war veteran. By doing so, he 'fails to understand that emotion is not a biological phenomenon which can be assessed objectively regardless of the circumstances, but a socially constructed entity specific to a particular context'. On that basis, it was correct that he was criticised as to the inappropriateness of comparing people trained to deal with fear with the reasonable person.[144] Such a standard is set too high;[145] is unachievable;[146] and requires people to act contrary to their natural instinct of self-preservation.[147]

Thirdly – and perhaps this is the most logical argument – as the rationale of the duress defence is that it seeks to excuse a perpetrator who gave in to a threat to which others of reasonable fortitude would have done likewise,[148] it must be asked how it can be just to punish a defendant for doing precisely what the reasonable person would have done? Simester and Sullivan have rightly observed that 'there will be occasions when we would expect even a person of reasonable firmness to be coerced into participating in murder'[149] and, as the Law Commission noted in its 2006 Report, '[l]ittle, if any, blame may attach to someone's decision to take part in a killing under duress'.[150] Indeed, the Law Commission quoted Elias J in the case of *R v Hasan* where he stated that defendants acting under duress 'are being punished for giving way to what will often be enormous fear and wholly understandable human frailty'.[151] Indeed, in the same case in the House of Lords, Lord Bingham said that the argument for extending the defence to murder was logically 'irresistible', while Baroness Hale acknowledged both sides of the duress coin

140 Defined as an ordinary citizen; 'a hypothetical person in society who exercises average care, skill, and judgment in conduct', *Free Dictionary* <http://legal-dictionary.thefreedictionary.com/Reasonable+Person>. For more on the reasonable person, see for example, G P Fletcher, *Rethinking Criminal Law* (Little, Brown & Co, 1978) 247; V Nourse, 'After the Reasonable Man: Getting over the Subjectivity/Objectivity Question' (2008) 11 New Criminal Law Review 33; and P Westen, 'Individualizing the Reasonable Person in Criminal Law' (2008) 2 Criminal Law and Philosophy 137.

141 This is a claim made by both R A Duff, 'Virtue, Vice and Criminal Liability: Do We Want an Aristotelian Criminal Law?' (2002-3) 6 Buff Criminal Law Review 147 his n 48 and D Pascoe, Thesis (ANU) 'Murder and the Defence of Necessity' (2007) 16 referencing Yeo in 'Necessity under the Griffith Code and the Common Law (1991) 15 Criminal Law Journal 17, 36.

142 *Per* Lord Coleridge in *R v Dudley and Stephens* (1884) 14 QBD 273 at 287. Criticised by Norrie (n 81) 157.

143 *R v Howe* [1987] 1 AC 417, 432. See criticism of this by E Colvin 'Ordinary and Reasonable People: The Design of Objective Tests of Criminal Responsibility' (2001) 27 Monash University Law Review 197, 214.

144 Reilly (n 10) 148 and 149 and see also J Gardner 'Justifications and Reasons' in A P Simester and A T H Smith (eds), *Harm and Culpability* (Clarendon Press 1996) 121; and the special 'Courage' issue (n 100).

145 A Noti, 'The uplifted knife: morality, justification and the choice-of-evils doctrine' (2003) 78 New York University Law Review 1859, 1886.

146 F Leverick, 'Defending self-defence' (2007) 27 Oxford Journal of Legal Studies 563, 570; Noti (n 145) 1886.

147 M Kremnitzer, 'Proportionality and the psychotic aggressor: another view' (1983) 18 Israel Law Review 178, 201.

148 S M H Yeo, 'Proportionality in criminal defences' (1988) 12 Criminal Law Journal 211, 219; S H Kadish, 'Excusing crime' (1987) 75 California Law Review 257; and A Brudner, 'A theory of necessity' (1987) 7 (3) Oxford Journal of Legal Studies 339 are just three who have made this point.

149 And went on to say that excluding murder and attempted murder from the duress defence is 'too rigid'; *Simester and Sullivan* (n 17) 805.

150 Law Commission (n 12) [1.54].

151 Ibid [6.51].

when she said, on the one hand, that she 'did not understand why the defendant's beliefs and personal characteristics are not morally relevant to whether she could reasonably have been expected to resist' while, on the other, she accepted 'that even the person with a knife at her back has a choice of whether or not to do as the knifeman says'.[152]

In these brief quotations, Bingham and Hale identify three of the problematic issues in the duress defence:

(1) extending the defence to murder;

(2) that characteristics are relevant to the capacity to resist; and

(3) that the duressee has a choice of whether to comply with the threat or not.

(1) Baroness Hale picked up on Lord Bingham's comment on extending the duress defence when she noted that the Law Commission had 'sold out to subjectivism' when it recommended in its 2006 Report on *Murder, Manslaughter and Infanticide* that duress 'should be a full defence to … murder and attempted murder'. At that time, the Law Commission had said it was 'wrong even in respect of murder to condemn the defendant for not acting heroically rather than reasonably'.[153]

In essence, there are two separate points here – subjectivism – and the argument for a full defence. The latter is not the main focus of this article, but one possible bone of contention with the latter is that permitting duress as a full defence to murder would lead to an acquittal, and not, as is the case in the partial defence of loss of control, to a conviction for manslaughter. This would then seemingly give fear – which goes to the very heart of duress[154] – a superior excusatory authority than anger (and fear) in loss of control.[155]

Subjectivisation of course, goes to the heart of the objective test, the purpose of which, it has been said, is to specifically 'impose a test of courage':[156] i.e. the reasonable person is created to epitomise a typical member of society;[157] he/she

> embodies a reasonable level of courage demanded of *all* citizens [and as such] makes no concessions to those who are more 'pliant' or susceptible to threats … [There is a] universal expectation … that the law is to be observed. [As Lord Lane said:] it was 'a matter of public policy … to limit the defence … by means of an objective criterion formulated in terms of reasonableness'.[158]

This means that the law deals with everyone in the same way, even though we are not the same.[159] Some of the reasons for this are that the objective standard 'is *intended* to hold persons up to a desirable standard of conduct' in order to improve their behaviour which, as a consequence, also benefits society; that the law cannot take into account all of the quirks from which people suffer; and that it would be hard pressed for the legal system to determine the exact capabilities of every person.[160] The undeniable contrary argument

152 *R v Hasan* [2005] 2 WLR 709 *per* Lord Bingham [21] and *per* Baroness Hale [73].

153 Law Commission (n 12) [6.21].

154 Loveless (n 139) 663.

155 Reilly (n 10) 146.

156 Virgo (n 128).

157 L Dahan-Katz, 'The implications of heuristics and biases. research on moral and legal responsibility' in Vincent (n 46) 149.

158 Smith (n 129) 370.

159 C Bublitz and R Merkel, 'Guilty minds in washed brains' in Vincent (n 46) 366. See also N A Vincent, 'Enhancing responsibility' in Vincent (n 46) 318.

160 Dahan-Katz (n 157) 149 and 150.

to this is that if the law ascribed all of the defendant's characteristic to the reasonable person, then there would be no point to having an objective test, as it would be totally negated.[161]

An obvious criticism of refusing to admit characteristics is that it is unfair for a person who may be unusually susceptible and who is simply incapable of resisting the threats to be judged according to the objective standard.[162] Virgo goes on to say that this 'misses the key point. Such a characteristic is not relevant because it contradicts the rationale of the objective test which seeks to determine the boundaries of courage.'[163] As was noted above, this is in fact a standard of heroism which is very often unachievable even by the hypothetical reasonable person.[164]

(2) Virgo's comment refers to defendants who are incapable of resisting the threats. This leads to the second point as expressed by Baroness Hale above – that she did not understand why the defendant's beliefs and personal characteristics are not morally relevant to whether she could reasonably have been expected to resist. The issue with this is that – as will be seen in the neuroscience section below – it is not possible to ascertain whether the perpetrator was indeed not capable of resisting the threat because he was so fear-struck that he/she could not see any choice other than to comply[165] or whether he *was* capable, but chose not to exercise that capacity at the time. If the former, Tadros has argued that where the perpetrator is 'so overwhelmed by fear' and it is clear that he could not achieve a higher standard of courage than he did, then the duress defence should be available. Contrarily, '[i]f an agent has the capacity for courage … but chooses not to act courageously on this occasion … then the defendant has no excuse'.[166] It should be noted at this stage that, although neuroscience cannot provide an answer to this conundrum, it is nonetheless known that a person's reaction to fear operates at a conscious and subconscious level.[167] Put together with the fact that fear is subjective and dependent on the situation, and that its effects will differ as between every individual, it is still documented that 'the ability to control [the fear reaction] is limited'.[168]

(3) Tadros refers to the perpetrators 'choice', as Baroness Hale does when she states that 'even the person with a knife at her back has a choice of whether or not to do as the knifeman says'.

This takes us logically to the difficult third point above – that the duressee has a choice of whether to comply with the threat or not – and to a very brief exposition of choice and character theory which are at odds with each other in the context of duress and, to an extent, loss of control also, but which are also relevant to the element of control present in both defences.

As was noted in the introduction, both duress and loss of control are excusatory, rather than justificatory. Traditionally, justification focuses on what the perpetrator did, his choice of action, without taking into account any personal characteristics he

161 A slightly different version of this can be seen in Smith (n 129) 366.

162 Virgo (n 128).

163 Ibid.

164 See text to (n 141) and the Law Commission, text to (n 153) above.

165 Reilly (n 10) 146. What Fischer and Ravizza call 'irresistible fear', J M Fischer and M Ravizza, *Responsibility and Control: A Theory of Moral Responsibility* (Cambridge University Press 1998) 82.

166 Tadros (n 106) 314–315.

167 See (n 46) above.

168 Claydon and Rödiger (n 119), referring to Damasio and Le Doux's research.

possesses, while excuse focuses on the perpetrator's characteristics:[169] i.e. the former is compatible with the choice theory, where individuals are seen as being responsible for their conduct if it is freely chosen, but where such conduct would be excused where it is not.[170] Contrarily, character theory looks to features of the perpetrator's character, the 'fixed or stable aspects of an agent's psychological make-up' and not on whether his behaviour followed the normative paradigm.[171] It would seem therefore that loss of control and duress fit into character theory, as the test for both involves ascription of character, albeit to a limited extent. However, academic discourse has shown that, strangely, choice theory may be appropriate to the duress defence because the perpetrator *does* have a choice, albeit it is a restricted one.[172] If we return to Loveless' exposition on domestic violence and duress, she points out that:

> Staying with a violent partner … should not imply freedom of choice or autonomy. Her choices are determined by her duressor … The very real risk is that her response to violence for which she bears no responsibility is likely to be regarded as unreasonable whatever she does. What is socially acceptable 'fortitude' in a relationship where violence is endemic and choices curtailed?[173]

Loveless goes on to say that the victim's choices are constrained because she is 'controlled by the duressor',[174] and it is this notion of control which links duress, and indeed loss of control, with the ascription of responsibility because both defences are based on the principle that no one should be criminally liable for crimes committed as a consequence of influences outside of their control.[175]

In law, it is recognised that responsibility relies upon the fact that the perpetrator acts voluntarily and has the freedom or control to choose a course of action: i.e. there is an alternative course of action that is available to him and that he is free to make a choice as to which course of action he pursues.[176] In this, he needs to be acting voluntarily because voluntariness is a prerequisite of responsibility. In duress the duressee is not acting voluntarily, nor does he have the genuine freedom of choosing a course of action because, in fear of the imposed threat, he is controlled by the duressor.[177] Even if he is responsible, he should not be found to be culpable because he was acting under duress.[178] Similarly, in loss of control, the role of the partial defence reflects the fact that the perpetrator does not completely meet the requirements needed to be held responsible, and this includes his capacity to control his behaviour.[179]

---

169 B McSherry 'Criminal responsibility, "fleeting" states of mental impairment, and the power of self-control' (2004) 27 International Journal of Law and Psychiatry 445, 453.

170 C Finkelstein, 'Excuses and dispositions in criminal law' (2002) 6 Buffalo Criminal Law Review special issue 'The New Culpability: Motive, Character and Emotion in Criminal Law' 317, 319.

171 Ibid 324. The capacity theory may also be relevant in loss of control insofar as it refers to a person 'who lacks the capacity to restrain his emotions' 330.

172 *Simester and Sullivan* (n 17) 812. See also Smith's choice account of duress in Smith (n 129).

173 Loveless (n 119) 100.

174 Ibid 96.

175 Arenson (n 13) 67, although his focus is on duress and necessity alone. Or, as Levy and Bayne have written, '[c]ontrol is a necessary condition for moral responsibility'; Levy and Bayne (n 104) 465. Or as stated by M Pardo and D Patterson in *Minds, Brains and Law: The Conceptual Foundations of Law and Neuroscience* (Oxford University Press 2013) 37: 'A brain must be "in control" for a human being to make choices. That cannot be disputed.' Or see Fischer and Ravizza (n 165) 17.

176 Fischer and Ravizza (n 165) 20.

177 Loveless (n 139) 663: 'If the threat is, or is believed to be, violent, control will be maintained by fear …'.

178 Fischer and Ravizza (n 165) 83.

179 K Sifferd, 'Translating scientific evidence into the language of the "folk"' in Vincent (n 46) 186.

However, it is not as straightforward as this because, although in both duress and loss of control, the defendant is acting in circumstances engendering significant stress and powerful emotions,[180] in duress, the defendant is expected to try and resist his fear[181] and choose death, while, in loss of control (and in provocation, as it then was), where the defendant is angry, he is seen to be *unable* to choose. This has been interpreted as evidence of the disparity in the degree of control expected of a person who is acting out of fear as opposed to one acting out of anger.[182] Thus, as Dressler has elaborated:

> With duress, only Actor's choice-opportunities are reduced. As such, we demand that the unlucky Actor accept his unenviable choices, and make the morally 'right' decision … he is capable of making such a decision.
>
> In provocation cases, however … Our common experience informs us that anger affects choice-capabilities, not mere opportunities. Anger makes us less *able* to respond in [an] appropriate fashion.[183]

So, the argument is that if the defendant is angry, he is considered to be *incapable* of making a choice, but if he is afraid, he is considered to be *capable* of making a choice. Reilly, points to one of Lord Hailsham's *dicta* in *Howe*, in which the latter also highlights the division 'between anger under provocation *which is based in affect*, and fear under duress *which is based in cognition*'. As Reilly notes, the consequence of this is that Lord Hailsham recognises that a duressee's cognitive capacity to resist killing is greater in a situation of duress than it is in what was provocation[184] (now loss of control). For Lord Hailsham, the clue for the distinction lies in the difference between *affect* and *cognition*. As was seen earlier, the cognitive view of emotion holds that the actor can assess the importance or significance of events and can himself evaluate the appropriateness of the emotion. This includes fear, but also – according to Kahan and Nussbaum – anger. However, Lord Hailsham places anger into the noncognitive theory of emotion[185] and to the affectivist approach, i.e. that anger in loss of control 'is the result of uncontrollable, biological, physiological and neurological changes within the actor'.[186] Finkelstein agrees with this approach, saying that loss of control cases 'at least involve an altered psychological condition' and that, as such, the partial defence fits into the capacity theory.[187] It will be recalled that this refers to a perpetrator in loss of control who 'lacks the capacity to restrain his emotions'.[188] Spain, in the same vein, also argues that loss of control is 'linked to a mechanistic understanding of emotion'; that is, that anger in particular, is both uncontrollable and not open to evaluation.[189] On the other side, and agreeing with Kahan

180 Yeo (n 148) 211.

181 Uniacke (n 28) 111.

182 See, for example, Reilly (n 10).

183 Dressler (n 12) 463–464. Emphasis in original.

184 Reilly (n 10) 147. Emphasis added.

185 'Non-cognitive theories are those that defend the claim that judgments or appraisals are not part of the emotion process. Hence, the disagreement between the cognitive and the non-cognitive positions primarily entails the early part of the emotion process. The concern is what intervenes between the perception of a stimulus and the emotion response. The non-cognitive position is that the emotion response directly follows the perception of a relevant stimulus. Thus, instead of any sort of evaluation or judgment about the stimulus, the early part of the emotion process is thought to be reflex-like.': 'Theories of emotion', *Internet Encyclopedia of Philosophy* <www.iep.utm.edu/emotion/#SH4b>. This would seem to be akin to the mechanistic theory.

186 Reilly (n 10) 131.

187 Finkelstein (n 170) 334.

188 Ibid 330.

189 Spain (n 13) 17.

and Nussbaum, the Law Commission stated that '[t]he emotion aroused in the provoked contains a *cognitive* component, viz. the belief that the provoked has been wronged by the provoker'.[190]

It would seem then that dividing the two emotions in *this* way does not provide a definitive answer as to why more is expected from the fearful perpetrator in duress than from the angered perpetrator in loss of control. However, perhaps a second *dictum* from Lord Hailsham in *Howe*, will assist. In addition to the distinction he highlighted between affect and cognition noted above, he also went on to say:

> [p]rovocation … is a concession to human frailty due to the extent that even a reasonable man may, under sufficient provocation temporarily lose his self control [sic] towards the person who has provoked him enough. Duress … is a concession to human frailty in that it allows a reasonable man to make a conscious choice between the reality of the immediate threat and what he may reasonably regard as the lesser of two evils.[191]

This *dictum* recognises that both defences are a concession to human frailty, but that the basis for that concession is different. In provocation/loss of control it is because it is accepted that a reasonable man *can* lose his self-control. This is simply a given without any further explanation and goes to the question raised earlier in the duress section as to the justness of punishing a defendant for doing what the reasonable person would have done. The difference is, of course that he has a partial defence to murder in loss of control, whereas no defence at all is available to the duressee in murder, but if there was, it would result in a complete acquittal.

In contrast and according to Lord Hailsham, the basis of the concession to human frailty in duress is that it allows a reasonable man to make a choice between two evils, one involving his breaking the law and the other some evil to himself or others. The balance between two evils is traditionally perceived as a justificatory notion and has, in recent years, been used to justify 'killing' in the context of necessity,[192] most notably *Re A (Children) (Conjoined Twins)*.[193] As a rule, however, killing another is not considered to ever be the lesser of two evils,[194] and it must be assumed that this is what Lord Hailsham was rather obtusely referring to. Furthermore, and as has been seen, duress is an excuse and not a justification. Secondly, the criteria for duress does not require a balancing of two evils. Thirdly, the court in Howe relied on *R v Dudley and Stephens* as authority for holding that necessity/duress was not a defence to murder, but *Dudley and Stephens* was certainly not decided on the basis of the lesser of two evils; if it had been, the defendants would not have been convicted.

Having said all that, it should be emphasised that Lord Hailsham's *dictum* regarding the different bases of the concession to human frailty does not in point of fact explain why the perpetrator in duress has a choice, but the perpetrator in loss of control does not. The only possible answer lies in a statement made earlier that the 'dimension of control' is lower when angry than it is for any other emotion, including fear.[195] Indeed, Holton and

190 Law Commission (n 52) [3.39]. Emphasis added.

191 *R v Howe* [1987] AC 417; see Arenson's very interesting discourse on this (n 13) 70–71.

192 It has been argued that necessity and duress of circumstances are one and the same. They are not. See, for example, Commentary on *R v Jones (Margaret)* [2005] QB 259 (CA) by D Ormerod in [2005] Criminal Law Review 122.

193 *Re A (Children) (Conjoined Twins: Surgical Separation)* [2000] 4 All ER 961.

194 This goes to proportionality, an element of justificatory necessity.

195 See text to (n 67) above.

Shute, writing before the changes to provocation were implemented in the Coroners and Justice Act, wrote that anger *extinguishes* self-control.[196] The reason why fear is not perceived in the same way is because very often people's responses to fear operate at the subconscious level so that there is 'less *conscious* control of their behaviour' when fearful.[197] When this happens, the reaction is instinctive and tends not to involve a 'conscious thought process'.[198] Although patently treated differently in legal terms, Claydon and Rödiger suggest that responses to both these emotions are 'almost indistinguishable' and that both evidence from medicine and neuroscience suggests that 'there is no scientifically objective measure for the effective separation of anger and fear'.[199] This is not borne out by the descriptions of the two emotions set out earlier.

That neuroscientific study cannot conclusively show changes to the brain which would explain the way in which anger and fear are differently perceived in the law is unfortunate, but as a relatively recent literature review has shown,[200] other than a brief discourse on capacity to resist, neuroscientific studies have tended to focus more generally on responsibility and determinism. However, that limitation aside, as a relatively new scientific method which *could* provide evidence of whether a legitimate excuse existed or not,[201] neuroscience is relevant here because it can ascertain whether a person has the capacity to make a decision and, consequently, if that person can or should be held responsible.[202] It can also 'suggest new ways of looking at both the content and application of our concepts',[203] especially relevant where emotional mental states can lead to an understanding of the reasons why a person acted as she did: that is, people's behaviour 'cannot be understood if [mental states] are excluded'.[204]

4 The potential impact of neuroscientific study

In the 1990s the term 'neurolaw' was devised as a way of describing the links between law and neuroscience.[205] Similarly, the term 'neurolegalist' has been used to describe researchers who claim that 'scientific data about the brain can illuminate or transform law and legal issues'.[206]

Neuroscientific studies, the relevance of which have, quite rightly, become more evident in law during the last 15 years or so,[207] focus to a large extent on the way in which brain functions determine the way people behave and react and on whether it can be

196 Holton and Shute (n 80).

197 Le Doux (n 46). Emphasis added.

198 Claydon and Rödiger (n 119).

199 Ibid.

200 F X Shen, 'Law and neuroscience' (2016) 48 Arizona State Law Journal 1043.

201 S J Morse and W T Newsome, 'Criminal responsibility, criminal competence, and prediction of criminal behavior' in S J Morse and A L Roskies (eds), *A Primer on Criminal Law and Neuroscience* (Oxford University Press 2013) 153 and 155.

202 See for example, Sifferd (n 179) 185; and Pardo and Patterson (n 175) 134.

203 A R Mackor, 'What can neurosciences say about responsibility?' in Vincent (n 46) 76.

204 Morse and Newsome (n 201) 152.

205 J Chandler, 'Neurolaw and neuroethics' (2018) 27 Cambridge Quarterly of Healthcare Ethics 590 'Conclusion'. See text to (n 19).

206 Pardo and Patterson (n 175) xxvi.

207 See generally M Freeman (ed), *Law and Neuroscience. Current Legal Issue*s vol 13 (Oxford University Press 2011); S Zeki and O Goodenough, *Law and the Brain* (Oxford University Press 2006); B Garland (ed), *Neuroscience and the Law: Brain, Mind and the Scales of Justice* (Dana Press 2004); M Freeman and O R Goodenough (eds), *Law, Mind and Brain* (Ashgate 2009); and N A Farahany (ed), *The Impact of Behavioral Sciences on Criminal Law* (Oxford University Press 2009).

conclusively proved that our behaviour is predetermined. An examination of the neuroscientific literature, however, shows a clear divide between those who, on the one hand, advocate determinism and who argue 'that free will is an illusion', that no one is really answerable for his actions[208] and that the criminal law should be changed accordingly.[209] On the other hand, there are those who 'hold that given *current* scientific knowledge, there is no need to revise the law and its basic concepts'[210] and because, in any event, '[i]ncreased knowledge about brain chemistry will not answer normative questions about responsibility'.[211]

This then is the first domain where neuroscience could potentially be a useful tool, i.e. to assist in the ascription of responsibility.[212] Another issue is the 'outstanding problem'[213] of whether neuroscience can tell us whether a person simply did not have the necessary tolerance or restraint (loss of control) or courage (duress), or whether he did but chose not to exercise it.[214] A third – linked – issue is one of timing. Looking more closely at the current state of neuroscience and its examination of the effect of characteristic-type emotions on the brain, it would appear that it cannot resolve any of these issues.

As far as the brain is concerned then, the pre-frontal cortex (PFC) is 'considered [to be] the controlling mechanism [which keeps] our emotions in check'.[215] It 'regulates the inhibition of impulses'[216] by sending messages to the limbic system, especially the amygdala.[217] The latter is the section of the brain which governs emotions such as fear and anger.[218] Numerous studies show that damage to the PFC might make it hard for a person to control his conduct[219] and that this, in turn, could lead not only to more impulsive, but also to more violent, behaviour[220] but there is nothing to indicate that this is as a result of an *inability* to exercise self-control or simply because of a choice by the actor not to exercise self-control.[221] However, this is not definitive,[222] possibly because some of the studies have been relatively small.

---

208 Meynen (n 20) 820.

209 For example, Sapolsky, Greene and Cohen, and Davies. As such, it could be potentially 'destabilizing'; Abrams and Keren (n 11) 2026, and see the scenarios propounded by R M Sapolsky, 'The frontal cortex and the criminal justice system', and J Greene and J Cohen, 'For the law, neuroscience changes nothing and everything' both in Zeki and Goodenough (n 207).

210 Meynen (n 20) 820. Emphasis added. Such as Morse, Pardo and Patterson, and Michael S Moore, best known for *Causation and Responsibility* (Oxford University Press 2009).

211 Pardo and Patterson (n 175) 39. Similarly, Mackor (n 203) 57 and Shen (n 200).

212 See, for example, Sifferd (n 179) 184: 'Neuroscience provides substantial new and valuable information for determining criminal responsibility.'

213 Levy and Bayne (n 104) 468.

214 See, in relation to duress text to (nn 165–167) above.

215 B J Grey, 'Neuroscience and emotional harm in tort' in Freeman (n 207) 214.

216 W Glannon, 'What neuroscience can (and cannot) tell us about criminal responsibility' in Freeman (n 207) 17.

217 The 'emotion centre' Grey (n 215) 214. See in particular Le Doux (n 58) 288.

218 Glannon (n 216) 17; and Sapolsky (n 209) 235–238. As noted by Svendsen (n 26) 26 'people with damage to the amygdala are unable to feel fear'.

219 Glannon (n 216) 20.

220 A L Roskies and W Sinnett-Armstrong, 'Brain images as evidence in the criminal law' and M Freeman, 'Introduction: law and the brain' in Freeman (n 207) 112 and 4–5 respectively.

221 Freeman (n 220) 6. Although see Pardo and Patterson (n 175) 123–124 on this point.

222 Roskies and Sinnett-Armstrong (n 220) 112.

Results of neuroscientific studies to date suggest that it is also not possible to demonstrate that all actions are predetermined.[223] There are essentially four reasons for this. The first is because the majority of studies demonstrate only 'correlations rather than causation'.[224] At the moment, neuroscience simply does not have a way of determining whether a person 'lacked the capacity to refrain from performing a criminal act or had this capacity but failed to exercise it'.[225] Therefore, looking to the 'brain alone does not provide a satisfactory explanation of our actions',[226] and there is no conclusive evidence that the brain governs or causes behaviour. In fact, Pardo and Patterson go further than this and opine that 'neuroscience [is limited in that it] cannot tell us where the brain thinks, believes, knows, intends, or makes decisions. People (not brains) think, believe, know, intend and make decisions.'[227]

Linked with this, the second reason is that, irrespective of its cause, any weakness in the capacity to control behaviour does not in itself excuse or mitigate that behaviour. This is because the law is only interested in the actual mental state, and not what caused it.[228]

Thirdly, it must be remembered that behaviour is subject to the self-evident fact that each individual will respond in a different way to the same emotional stimuli. One reason for this is that emotions are not just natural biological phenomena; rather, they are influenced by the way we have been brought up and educated and by our experiences; social norms; environment; culture; chance; genetics, heredity[229] and other outside factors.[230] This is why it is so problematic to reconcile emotional reactions with the behaviour of the reasonable person and to comply with both the objective test[231] and the norms associated with criminal responsibility. Because our brains are able to adapt to external influences we are not constrained by one set of predetermined outcomes.[232] If this was not so, then no one would be responsible for anything.

The final reason then relates to the notion of legal responsibility. This is a normative paradigm constructed by society, whereby a person's conduct is measured against this hypothetical person of the defendant's sex and age (loss of control) or the sober person of reasonable firmness (duress).[233] It has been said that

> Neuroscience will never find the brain correlate of responsibility, because this is something we ascribe to humans, not to brains … psychiatrists … might be able to tell us what someone's mental state … is without being able to tell us … when someone has too little control to be held responsible. The issue of responsibility

---

223 E Aharoni et al, 'Can neurological evidence help courts assess criminal responsibility? lessons from law and neuroscience' (2008) 1124 Annals of the New York Academy of Sciences 145, 147.

224 Ibid and A L Roskies 'Brain imaging techniques' in Morse and Roskies (n 201) 68.

225 Glannon (n 19) 2.

226 Ibid 2.

227 Pardo and Patterson (n 175) 46.

228 S J Morse, 'Lost in translation? An essay on law and neuroscience' in Freeman (n 207) 531.

229 See, for example, Rosebury (n 70) 39; and C E Izard and E A Youngstrom, 'The activation and regulation of fear and anxiety' in Hope (n 59) 31.

230 See N A Farahany and J E Coleman JR, 'Genetics, neuroscience and criminal responsibility' and B Garland and M S Frankel 'Considering convergence: a policy dialogue about behavioural genetics, neuroscience and law' in Farahany (n 207).

231 A Reilly 'Loss of self-control in provocation' (1997-8) 21(6) Criminal Law Journal 320, 326.

232 M Belcher and A L Roskies, 'Neuroscience basics' in Morse and Roskies (n 201) 34. See also both Vincent in Vincent (n 159) 4 and S J Morse, 'Common criminal law and compatibilism' in Vincent (n 46) 39.

233 Farahany and Coleman (n 230) 240.

> … is a social choice … the idea of responsibility is a social construct and exists in the rules of the society.[234]

The final unresolvable issue relates to the 'time machine' problem.[235] In other words, neuroscience is unable to tell us what a defendant's state of mind was at the time the offence was committed because '[n]euroimaging can take place only after a crime has occurred. It is impossible to study the defendant's brain state at the time of the actual crime (in the past). Therefore, what actually happened at the moment of the crime in a defendant's brain is not directly accessible to neuroimaging findings after the crime.'[236] Put another way, '[t]oday's brain is not yesterday's brain'.[237]

Thus, although neuroscience has contributed to the study of emotion by providing new avenues of investigation[238] and different inroads into legal principles and practice, any such inroads will be both rare and uncertain for the conceivable future.[239] As Aharoni concedes, '[n]euroscience is much more limited in the kinds of conclusions it can support than the public, the legal system, and many neuroscientists would like to acknowledge'.[240] Thus, although academics such as Mitchell and Mackay have long argued that more 'effort should … be made … to have regard to what psychologists and psychiatrists have discovered about the ways in which emotions affect our mental states and our behaviour',[241] disappointingly, and currently at least, it seems that neuroscience is unable to convey any information that will contest the law as it stands;[242] rather, it is 'simply the most recent candidate … for explaining behavior deterministically … it adds nothing new'.[243]

Conceivably, however, in time, neuroscience might develop sufficiently to be able to discover more about a perpetrator's ability to resist, such that there would be evidence to support the use of the defences where an individual can truly show that his characteristics prevented him from being able to resist.[244]

234 M S Gazzaniga and M S Steven, 'Free will in the twenty first century: a discussion of neuroscience and the law' in Garland (n 207) 68. Likewise, Aharoni (n 223) 145–146.

235 H T Greely, 'Mind reading, neuroscience and the law' in Morse and Roskies (n 201) 120.

236 G Meynen, 'A neurolaw perspective on psychiatric assessments of criminal responsibility: decision-making, mental disorder, and the brain' (2013) 36 (2) International Journal of Law and Psychiatry 93, 96–97.

237 Roskies (n 224) 70. Timing would not, of course, be so relevant where the perpetrator suffered from a permanent mental disorder. A broader and different question would be whether such a person should be held criminally responsible more generally. I am grateful to the reviewer who raised this point. Notably, most of the research carried out has been on persons with permanent, as opposed to, transient, dysfunctions. See, for example, A L Glenn and A Raine, 'Ethical issues' in A L Glenn, *Psychopathy: An Introduction to Biological Findings and their Implications* (New York University Press 2016); Morse and Roskies (n 201) 247; Levy and Bayne (n 104); Nestor (n 11) and Meynen (n 236).

238 Reilly (n 10) 126.

239 Morse (n 228) 529.

240 Aharoni (n 223) 158.

241 Mackay and Mitchell (n 46) 49. In contrast, see Law Commission (n 52) [3.28].

242 Greene and Cohen (n 209) 224.

243 Morse and Newsome (n 201) 153.

244 A L Roskies and S J Morse, 'Neuroscience and the law: looking forward' in Morse and Roskies (n 201) 247.

Conclusion

More generally, research into emotions, neuroscience and the law since the beginning of the century[245] has made a substantial contribution to the way we think about emotions and the law. For example, Sanger, Bandes and Grossi's research has focused on the role and impact of emotions in the law, in law-making and on people who work in the law.[246] Sanger lists incidents where emotions play a part in our daily activities and gives examples of where emotion has resulted in 'a spate of laws, which are not only motivated by emotion, but include as part of their scheme some emotive component'.[247]

More specifically, it has been seen that the law does recognise that emotions play an essential role in effective and rational decision-making[248] and that anger and fear, the focus of this piece, influence the defendant's reasons for acting[249] and affect their subsequent behaviour, in a different way. It is ironic then that, despite the pervasiveness of fear in these criminal defences, its role there remains 'ambivalent'[250] – we have seen that fear has been introduced into the loss of control partial defence, where it does not 'fit' and where, in any event, it is not given the same priority as anger. Contrarily, it has not been introduced into duress, a defence which does not even acknowledge the fear that a threat engenders, although it is evident that fear affects, and motivates, the way in which an offender responds to the threat, and where it would therefore 'fit': i.e. fear should specifically be referred to as part of the rationale for the duress defence.

As to loss of control, it has been seen here that anger has a different effect on an actor as compared to fear and, as such, the way in which fear has been accommodated in the loss of control partial defence does not do it justice. The partial defence, as its predecessor did before it,[251] still expects individuals to suffer a loss of control. However, although the partial defence is dependent upon a loss of control arising as a result of some emotional turmoil, the emotional turmoil giving rise to the loss of control is not its defining feature: that falls to the necessity to prove an actual loss of control. This does not lie easily with the manifestations of fear.

Had the loss of control requirement been abolished, as was recommended by the Law Commission,[252] there would perhaps have been a better opportunity to consider *emotion-based* alternatives as a basis for the defence, such as, for example the provisions on 'extraordinary emergency' referred to in section 25 of the Criminal Codes of Queensland and Western Australia (the Griffith Code),[253] or the notion of 'extreme mental or emotional disturbance' as referred to in clause 210.3.(1)(b) of the Model Penal Code (MPC). The latter provides that:

> [A] homicide which would otherwise be murder [is manslaughter when it] is committed under the influence of extreme mental or emotional disturbance for which there is reasonable explanation or excuse. The reasonableness of such

245 Nestor (n 11) 1.

246 See (n 11) above.

247 e.g. Megan's law; Sanger (n 11) 110. See also Abrams and Keren (n 11) 2060.

248 Damasio (n 30) 40, and Abrams and Keren (n 11) 2004.

249 Tadros (n 106) 305.

250 Reilly (n 10) 121.

251 *Per* Lord Hoffman in *R v Smith (Morgan)* [2001] 1 AC 146, 173, noted in Law Commission (n 81) [1.40], and *Lowe* [2003] EWCA Crim 677, noted in Law Commission (n 52) [3.117].

252 Law Commission (n 52) [3.115].

253 See S M H Yeo, 'Necessity under the Griffith Code and the common law' (1991) 15 Criminal Law Journal 17, and Pascoe (n 141).

> explanation or excuse shall be determined from the viewpoint of a person in the actor's situation under the circumstances as he believes them to be.

The Law Commission considered this in its 2004 *Report on Partial Defences*, noting that while a number of US states ('reform states') had adopted the clause,[254] some had omitted the 'actor's situation' formulation contained in the second sentence. In line with that, Professor Sanford Kadish, in an interesting report written for the Law Commission, advised that the 'actor's situation' formulation should be rejected for a number of reasons, not least because it was not consistent with the rationale of the partial defence.[255] However, he commented that 'the EED formulation … is a felicitous and useful improvement over the traditional formulation … [which] hits the nail on the head in asking whether there is reasonable explanation or excuse for the mental disturbance, since in these cases the basis of the mitigation is precisely the emotional disturbance … this part of the MPC proposal has it right'.[256] Despite this endorsement, however, the Law Commission ultimately rejected the test as being unduly 'vague and indiscriminate', preferring instead that 'the defendant had legitimate ground to feel seriously wronged by the person at whom his or her conduct was aimed' as the basis of the defence.[257]

As to duress, Eimear Spain has staunchly advocated an emotion-based excusatory duress defence based on a 'reasonable emotional response' whereby the moral quality of the emotion and the individual's response to it could and should be evaluated.[258] According to Spain, the test could be that 'a reasonable person sharing the characteristics of the accused would have experienced fear in the circumstances', i.e. the jury would have to decide 'whether the fear experienced by the defendant (and subsequent action) was morally appropriate'.[259] The same test could be applied to loss of control.[260]

A further alternative incorporating *both* defences (due to their categorisation as excuses) has been advocated by Claire Finkelstein. She opts for what she calls 'rational excuses'. These 'straddle the line between traditional excuses and justifications' neither of which she claims are really appropriate. Rather,

> rational excuses exonerate because they are cases in which the agent acts on the basis of … an 'adaptive disposition,' namely a disposition that enhances the welfare of the agent who cultivates it and makes possible collective welfare improvements. The law is interested in promoting and rewarding such dispositions, because there will be mutual gains from their adoption if members of society generally possess them.[261]

One element of this is the recognition that characteristics *should* be taken into account, although Finkelstein's main point is that, unlike traditional character theory, she argues that the basis of permitting an excusatory defence is to excuse when the agent acts '*out of*

254 Law Commission (n 52) [3.50].

255 Ibid Appendix F [20–23].

256 Ibid [14] i.e. omitting the word 'mental' from the formulation, (preferring simply 'extreme emotional disturbance (EED)'. See Mitchell and MacKay above, text to (n 112).

257 Law Commission (n 52) [3.59].

258 Spain (n 13) 264, 268–270. Contrast Nourse who argues against an objective requirement of reasonableness; V Nourse, 'Passions's progress: modern law reform and the provocation defence' (1996–1997) Yale Law Journal 1331.

259 This would be excusatory and not a justificatory/lesser evils defence. Spain (n 13), 264 and see 272–273.

260 Albeit different considerations will apply where the fear element has been long-standing; Edwards (n 117) 228.

261 Finkelstein (n 170) 320–321.

character', rather than acting in character, because the former is 'most chosen', and reflects when he is at his most responsible.[262]

To sum up then, no preference is expressed here as to which option is favoured except to say that, firstly, it is recommended that in loss of control the fear emotion should be more properly incorporated into the defence such that it more effectively includes those whom it was intended to protect (such as, for example, victims of domestic abuse, but not limited to that class).[263] As it stands, the defence currently makes killing out of anger more defensible than killing out of fear.[264] This wrongly clings to the old (and now defunct) notion of honour from whence the anger derived and upon which the previous partial defence of provocation was based,[265] but it is also noted here that anger is not even specifically named as a trigger in the current defence definition.

Secondly, the duress defence should take into account the fear emotion, which is, after all, its key component. As it stands, the defence in its current form is too strict[266] in denying the defence to those who cannot help that they are not courageous or heroic enough. Perhaps it is time to put duress on an equal footing with loss of control, not least on the basis that a person who kills under duress is significantly less blameworthy than one who kills as a result of loss of control. [267]

The recommendations are thus founded on the arguments made here that fear is an appropriate emotion.[268] Both recommendations clearly advocate fear-centred reasoning as a basis for both defences. What part the objective reasonableness element of the defences should continue to play has been briefly questioned.[269] We know that the hypothetical reasonable person is a 'figure constructed to represent a standard member of society',[270] the purpose of whom is to ensure that everyone complies with societal norms, whether they have the capacity to do so or not.[271] Views on this rely on how committed we are to adhering to universal standards of conduct and/or being more open to the more subjective consideration of emotion as recommended here.

262 Ibid 343. Emphasis in original. This is very much like the view expressed by Tadros above at text to (n 106).

263 See, for example, Loveless (n 139) 662.

264 See, for example, Allen (n 15) 241–242.

265 See (n 9); (n 15); Dressler (n 12); Gough (n 42); and Allen (n 15).

266 *Simester and Sullivan* (n 17) 805.

267 Arenson (n 13). He claims that the difference between the two defences presents a 'gross imbalance'; it demonstrates 'an indefensible and strange dichotomy'; it is 'one of the most longstanding and obvious paradoxes in the criminal law' 77, 78 and 74 respectively.

268 Text to (nn 73–81).

269 Indirectly on page 343 for loss of control and more specifically for duress on pages 419, 420–421, especially text to (n 161).

270 Dahan-Katz (n 157) 149. See (n 140) and text to (nn 156–161) above.

271 Dahan-Katz (n 157) 149; Bublitz and Merkel (n 159) 362.

NILQ 71(3): 429–455

Recalibrating the governance of remedial clauses in contract law

MOSHOOD ABDUSSALAM

Auckland University of Technology, New Zealand

Abstract

Yawning gaps in bargaining powers between transacting parties have always been a source of concern in commercial relations and the legal governance of such relations. In modern times, the likely implications of gaps in bargaining powers are not only palpable as it concerns the affairs of transacting parties with weaker bargaining powers, but also on the welfare of society, at large. That is particularly so in this milieu of pervasive oligopolistic market structures, organised commercial networks, digitisation, and big data. The imperative to guard against the use of contractually agreed remedial clauses to consolidate market power and as tools for wealth extraction is the concern of this article. To this end, this article makes a case for a recalibration of the rule against penalties in contract law.

Keywords: contracts; liquidated damages; bargaining powers; wealth extraction.

> *Society suffers when any of the pillars weakens or strengthens overly relative to the others. Too weak the markets and society becomes unproductive, too weak a community and society tends toward crony capitalism, too weak the state and society turns fearful and apathetic. Conversely, too much market and society becomes inequitable, too much community and society becomes static, and too much state and society becomes authoritarian. A balance is essential!*
>
> Raghuram Rajan, *The Third Pillar* (William Collins 2020) xviii.

Introduction

One scholar who appears assertive in the call for judicial intervention in extortionate and oppressive bargains is Professor Stephan Waddams; who, in his recent book founded his case on equitable reasoning, chiefly on the axiom of equity that 'no man should be a gainer by another's loss'.[1] However, Waddams does not base his call on any discernible theoretical foundation beyond a firm reference to case law. This article seeks to carry his call further, from a different light. This paper deploys heterodox law and economics reasoning toward that end. The thesis of this article is that compensation for losses resulting from contractual wrongs should correspond to the loss suffered by a promisee, except in cases where the promisee can demonstrate legitimate interests deserving of super-compensatory protection. The thesis is pursued based on three strands, which are as follows:

1 Stephan Waddams, *Sanctity of Contracts in a Secular Age* (1st edn, Cambridge University Press 2019) 23–24, 129 and 232.

a. the courts should have wide powers to read down remedial clauses;

b. the imperative to reconceptualise all remedial clauses, including termination clauses, under one single category; and

c. the postulation and discussion of considerations that should guide courts in governing remedial clauses.

The first strand is the location of judicial power to revise extortionate and oppressive bargains in the advancement of social interests. In contract law scholarship, the significant interests commonly recognised as falling within the purview of contract law are performance, reliance and restitution interests.[2] The consideration is that contracts seek to promote mutually beneficial bargains between parties who have voluntarily assumed them. Unfortunately, however, the interests of the society within which contractual transactions occur are rarely factored into contract law rules, except generally in those rules prohibiting illegality. The assumption is that, so long as parties have mutually bound themselves to contracts, society would gain from the exchange of economic value between such parties.[3] Sadly, this is not always the case. Contracts that appear mutually beneficial to parties may yet impose or create the risk of a high cost to society.[4]

The second and third strands of this article focus on agreed remedies clauses, discussing them as species of contractual terms that are veritable candidates for judicial intervention; mainly on account of social interests. These strands advance the argument that all remedial clauses should be treated as one and should be judicially supervised or governed in uniform fashion. Through these strands, the gulf in the approaches of the UK Supreme Court and the Australian High Court towards agreed remedies as reflected in the cases of *Cavendish Square Holding BV v Talal El Makdessi*[5] and *Andrews v ANZ Banking Group Ltd*[6] is revisited. The thesis of this article concurs with the Australian approach as it is amenable to advancing social interests. However, despite the Australian approach being preferable to the UK approach, both share a similar flaw, and this is a concern echoed by Judge James Allsop of Australia, writing in extra-curial capacity.[7] As he observes, 'neither court … has taken the opportunity fully to fashion the doctrine for modern commerce in a nuanced way … to distinguish between contractual circumstances and to make evaluative judgments as to the presence or absence of oppression and extravagance informed by the equality or inequality of bargaining power and like considerations'.[8]

This article submits that an agenda towards fashioning a viable model for the governance of agreed remedies suitable to the needs of modern commerce requires that we pay close attention to prevailing socio-economic realities. The incidence of ever-increasing and widened gaps in bargaining powers in the marketplace is the bane of the unequal distribution of cooperative gains in commercial dealings. The feared severity of the situation is the reason why some leading economists have decried a 'rigged' economic

2 See, L L Fuller and W R Perdue Jr, 'The reliance interest in contract damages: 1' (1936) 46 Yale Law Journal 52; see also, the judgment of the Australian High Court in *Clark v Macourt* [2013] HCA 56.

3 See, Robert Cooter and Thomas Ulen, *Law and Economics* (6th edn, Pearson Education 2014) 336.

4 See, Eric Talley, 'Contract renegotiation, mechanism design, and the liquidated damages rule' (1994) 46 Stanford Law Review 1195; see also, Benjamin Hermalin, Avery W Katz and Richard Craswell, 'The law and economics of contracts' (2006) Handbook of Law and Economics 33–34.

5 [2015] UKSC 67.

6 (2012) 247 CLR 205.

7 James Allsop, 'Singapore Academy of Law Distinguished Speaker Lecture 2017: The doctrine of penalties in modern contract law' (2018) 30 Singapore Academy of Law Journal 1.

8 Ibid 26.

state characterised by uneven wealth distribution, economic inequality and, arguably, reduced economic growth.[9] In response to that position, some other economists have denied any such perilous state of affairs; arguing instead that, although there might be uneven wealth distribution, the living conditions of people across the world have improved significantly in comparison to previously prevailing states.[10] Regardless of whatever position one takes in that debate, one cannot deny that wide gaps in bargaining powers can have significant implications for contractual outcomes, particularly the distribution of cooperative gains between parties to an exchange. In short, in any marketplace, the party with strong bargaining powers is king!

The truth is that uneven bargaining powers do not only result in an uneven distribution of exchange gains. They can create serious social problems, particularly an attenuation in market competition, entrepreneurship, innovation and ultimately, social welfare decimation.[11] The risk of these unfavourable social states arising owing to gaps in bargaining power is palpable in the modern economic setting, which is shaped by a combination of pervasive oligopolistic market structures, organised commercial networks, digitisation and big data.[12] Scholars have argued that competition/anti-trust law policies appear inadequate in addressing these new commercial realities and challenges.[13] Thus, it is this current state of the marketplace that informs the call for judicial intervention in remedial clauses contrary to the opinion of some scholars that the market has not witnessed any change warranting such legal reformation.[14]

This article does not address consumer protection issues, as it focuses solely on commercial bargains; and its presentation takes the following steps. Section 1 discusses the place of agreed remedies in the scheme of contracts along with an examination of arguments against the penalty rule. Section 2, addressing the first strand of the article's thesis, examines the essence of contract law; making the submission that contract law as an institution should be recognised as one which aims at enabling private empowerment through exchanges. Also, that bargains that cross the lines of empowerment should be candidates for judicial scrutiny and review. Section 3, anchoring the second strand, advances the argument that all remedial clauses, regardless of form or label, are same in substance and, as such, should be uniformly governed. Section 4, dealing with the third strand, postulates and expounds on considerations that should govern the judicial governance of remedial clauses. In the final section, the article concludes by connecting the three strands together.

9 See, Brink Lindsey and Steven Michael Teles, *The Captured Economy: How the Powerful Enrich Themselves, Slow Down Growth, and Increase Inequality* (Oxford University Press 2017); see also, Joseph Stiglitz, *The Price of Inequality: How Today's Divided Society Endangers our Future* (1st edn, WW Norton & Co 2012); see also, Joseph Stiglitz, *People, Power, and Profits: Progressive Capitalism for an Age of Discontent* (1st edn, WW Norton & Co 2019).

10 See Jean-Philippe Delsol, *Nicolas Lecaussin and Emmanuel Martin (eds), Anti-Piketty: Capital for the 21st Century* (Cato Institute 2017).

11 See Jonathan Tepper and Denise Hearn, *The Myth of Capitalism: Monopolies and the Death of Competition* (1st edn, Wiley 2018); see also, Tim Wu, *The Curse of Bigness: Antitrust in the New Gilded Age* (1st edn, Columbia Global Reports 2018); see also, Robert Reich, *Saving Capitalism: For the Many, Not the Few* (1st edn, Vintage 2016)

12 See, Margaret Radin, 'The deformation of contract in the information society' (2017) 37 Oxford Journal of Legal Studies 505; see also, Robin Kar and Margaret Radin, 'Pseudo-contract and shared meaning analysis' (2019) 132 Harvard Law Review 1135; see also, Maurice Stucke and Allen Grunes, *Big Data and Competition Policy* (1st edn, Oxford University Press 2016).

13 See, Ariel Ezrachi and Maurice Stucke, *Virtual Competition: The Promise and Perils of the Algorithm-Driven Economy* (1st edn, Harvard University Press 2016);

14 John Carter, Wayne Courtney, Elisabeth Peden, Andrew Stewart and G J Tolhurst, 'Contractual penalties: resurrecting the equitable jurisdiction' (2013) 30 Journal of Contract Law 99, 113.

1 The place of agreed remedies in the scheme of contracts and the case against the penalty rule

When parties enter into contracts, they agree on terms that would govern their relations by adopting correlative rights and obligations. In some cases, they may agree on consequences that should follow upon specified events. Falling within this latter class of terms are those clauses that qualify as agreed remedies clauses. The law classifies such remedial clauses based on their form. These clauses, despite their different labels, have the same purpose(s), which is vesting the promisee with powers or rights to act in ways towards the enforcement of contracts. In other words, they are self-help remedies.[15] The major problem which besets both the judicial governance and legal analysis of these clauses is the adherence to categorisation. The traditional position, for example, distinguishes the jurisdiction to reform forfeiture clauses from that to invalidate penalty clauses.[16] Such bifurcation is impractical, especially as contractual drafting may make either category of provisions indistinguishable from the other. Judges have long recognised this difficulty. Lord Denning exposed the facileness of such distinction in a prominent case;[17] and in another case, in which he acknowledged such legal difference, he expressed himself as willing to disregard it.[18] Even in *Dunlop Pneumatic Tyre Co Ltd v New Garage Motor Co Ltd*,[19] which was, until recently, recognised as providing the modern restatement on agreed remedies, Lord Dunedin described the question as to which jurisdiction governs what type of clause as being 'probably more interesting than material'.[20]

At this juncture, we turn to discuss the arguments against judicial reformation of remedial terms in business contexts. Here, we highlight the perceived commercial purposes of remedial terms.

1.1 Arguments in favour of the strict enforcement of remedial clauses

Some scholars have called for the abandonment of judicial rules that seek to govern agreed remedies. Cases of this kind abound, largely concerning liquidated damages and forfeiture clauses. Proponents of this line of reasoning have pursued their arguments mostly on the need to respect individual choice and autonomy, on the ground that such provisions are as much a contractual term as those dealing with a contract price.[21] One such proponent is Sarah Worthington, whose argument for the abolition of the penalty rules was greeted with rejection by the UK Supreme Court in *Cavendish*.[22] The crux of her argument is that rules against penalties lack proper justification beyond moral outrage

---

15 See Celia Taylor, 'Self-help in contract law: an exploration and proposal' (1988) 33 Wake Forest Law Review 839; see also, Mark Gergen, 'A theory of self-help remedies in contract' (2009) 89 Boston University Law Review 1397.

16 See, Neil Andrews, Malcolm Clarke, Andrew Tettenborn, Graham Virgo, *Contractual Duties: Performance, Breach, Termination and Remedies* (2nd edn, Sweet & Maxwell/Thomson Reuters 2017) 523–559.

17 *Bridge v Campbell Discount Co Ltd* [1962] 2 WLR 439, 458–560.

18 *Stockloser v Johnson* [1954] 2 WLR 439, 450.

19 [1915] AC 79

20 Ibid 87.

21 See Mindy Chen-Wishart, 'Controlling the power to agree damages' in Peter Birks, (ed), *Wrongs and Remedies in the Twenty-First Century* (Oxford University Press 2006) 280; see also, Sarah Worthington, 'The death of penalties in two legal cultures?' (2016) 7 UK Supreme Court Yearbook 129, 140.

22 Sarah Worthington, 'Common law values: the role of party autonomy in private law' in Andrew Robertson and Michael Tilbury (eds), *The Common Law of Obligations* (Hart 2016) 301

entrenched in case law.[23] She goes on to state that this absence of justification is the reason why judicial scrutiny of such clauses has always been pursued using crude and selective methods.[24] However, she shows some sympathy for the judicial intervention in forfeiture clauses; and this is simply on the basis that relief from forfeiture is only available to a promisor who can adequately address the injury of the promisee commensurably with contractual expectations.[25]

The most systematic set of arguments against penalty rules comes from conservative scholars of law and economics. The consolidation of this brand of scholars brought about an onslaught against several legal rules and ideas, most notably competition law.[26] In this vein, the judicial rules for the governance of agreed damages also received criticism. Although most of these scholars in the formative years of the movement were from the USA, a few of them were from Commonwealth jurisdictions. One of the latter was Trebilcock, who argued that judicial refusal to enforce such clauses creates the sub-optimal risk of increasing the promisee's costs in seeking to ensure performance.[27]

A line of attack popular with the scholarly movement is that remedial terms serve a signalling function between contracting parties.[28] They enable promisees to communicate to would-be promisors what their idiosyncratic expectations are; and the would-be promisor is also able to signal their ability to shoulder such expectations. Some scholars take the view that remedial provisions are very useful in franchising contexts; especially that they protect promisee–franchisors against the risks of 'difficult-to-verify' injuries or non-pecuniary losses such as the denting of trademark or brand value.[29] However, others have equally used empirical evidence to show that the usefulness or success of such clauses in franchising contexts is mixed.[30] Their submission is that such provisions may disrupt the relational dynamics between franchisees and franchisors.[31]

Another line of attack is that the non-enforcement of such clauses may distort the efficient allocation of commercial risks. In this connection, by relieving against such provisions, a promisor may secure a windfall for themselves after having signalled themselves as being of a particular capability to gain acceptance from the promisee who may not otherwise have offered them a bargain.[32] Some other connected arguments are that the non-enforcement of such clauses would increase transaction and litigation costs,

23 Ibid 316.

24 Ibid 319

25 Ibid 321–322.

26 See Steve Teles, *The Rise of the Conservative Legal Movement: The Battle for Control of the Law* (1st edn, Princeton University Press 2008) 122–123.

27 See Michael Trebilcock, *The Limits of Freedom of Contract* (Harvard University Press 1997); see also, Michael Trebilcock, 'The doctrine of inequality of bargaining power: post-Benthamite economics in the House of Lords' (1976) 26 University of Toronto Law Journal 359.

28 See, David Haddock, Fred McChesney, and Menahem Spiegel, 'An ordinary economic rationale for extraordinary legal sanctions' (1990) 78 California Law Review 1.

29 See Albert Choi and George Triantis, 'Completing contracts in the shadow of costly verification' (2008) 37 Journal of Legal Studies 503; Benjamin B Reed, 'Liquidated damages provisions: strategic drafting and enforcement issues' (2018) 37 Franchise Law Journal 523.

30 See Adam Badawi, 'Relational governance and contract damages: evidence from franchising' (2010) 7 Journal of Empirical Legal Studies 743.

31 Ibid 744–745.

32 See Charles Goetz and Robert Scott, 'Liquidated damages, penalties and the just compensation principle: some notes on an enforcement model and a theory of efficient breach' (1977) 77 Columbia Law Review 554.

along with inefficiencies in the marketplace.[33] The upshot of this may cause essential facilities such as loans or credit facilities to be withheld from needy entities, or cause the price of such facilities to be hiked.[34]

Finally, another basis of attack is that judicial reform of such clauses is anachronistic to modern market conditions and, therefore, should only apply in cases of procedural unconscionability such as (unilateral) mistakes and duress.[35] In this regard, the argument goes that the ascendency of the compensatory principle in contract law is a by-product of historical accident and path dependency.[36] There are several other arguments against judicial governance of such clauses; but ,for the sake of space, this article only highlights these salient ones mentioned.

The case against judicial intervention, as one can decipher, is built on respecting private ordering so long as there is a balance of fairness between contracting parties. In other words, that unconscionability is the only defensible ground for the curtailment of such clauses. The objective of this article is to show that commutative fairness in contract law would be mythical or ineffectual without a well-calibrated distribution of legal entitlements between contracting parties. In other words, commutative justice in contractual contexts is contingent upon distributive justice.

Some scholars trenchantly hold the view that contract law's domain is to enable exchanges and ensure fairness in bargains, but that issues relating to wealth distribution are not concerns of contract law.[37] Such a position is unsustainable, as this article argues. Judicial intervention in agreed remedies clauses is apt to serve the much-needed distributive justice goals of counterbalancing the vast chasm in bargaining powers and the use of contracts as a mechanism for rent extraction and the consolidation of market power. Judge Posner in the US Court of Appeals case of *Lake River Corp v Carborundum Co*[38] acknowledged that agreed remedies clauses could have social cost implications, saying:

> On the other side it can be pointed out that by raising the cost of a breach of contract to the contract breaker, a penalty clause increases the risk to his other creditors; increases (what is the same thing and more, because bankruptcy imposes 'deadweight' social costs) the risk of bankruptcy; and could amplify the business cycle by increasing the number of bankruptcies in bad times, which is when contracts are most likely to be broken.[39]

In that case, Judge Posner found the clause in issue to be an unenforceable penalty. However, being of firm conservative mould, he went on to describe the likelihood of unfavourable social outcomes resulting from remedial clauses as though they were unavoidable bad weather by saying: 'But since little effort is made to prevent businessmen

33 See Aristides N Hatzis, 'Having the cake and eating it too: efficient penalty clauses in common and civil contract law' (2002) 22 International Review of Law and Economics 381; see also Ugo Mattei, 'The comparative law and economics of penalty clauses in contracts' (1995) 43 American Journal of Comparative Law 427.

34 Deborah Zalesne, 'Enforcing the contract at all (social) costs: the boundary between private contract law and the public interest' (2005) 11 Texas Wesleyan Law Review 579.

35 See Larry Dimatteo, 'A theory of efficient penalty: eliminating the law of liquidated damages American' (2008) 38 Business Law Journal 633.

36 Robert E Scott and George G, Triantis, 'Embedded options and the case against compensation in contract law' (2004) 104 Columbia Law Review 1428.

37 See James Gordley and Hao Jiang, 'Contract as voluntary commutative justice' Tulane Public Law Research Paper No 19-3 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3324001>.

38 769 F2d 1284 (1985).

39 Ibid 1289.

from assuming risks, these reasons are no better than makeweights.'[40] Statements like these represent the views of most conservative scholars. With heterodox law and economics reasoning, this article provides counterarguments against such conservative views.

2 The first strand: private empowerment as the essence of modern contract law

Towards the advancement of the thesis of this article, it is crucial to identify contract law as a mechanism for private empowerment. Before pursuing that argument, however, it is vital to deconstruct the marketplace. The objective here is to show that the market does not operate in *laissez-faire* fashion, in the 'natural' and self-executing sense, as some scholars are wont to argue. The market and its outcomes operate in the shadow of the law's distribution of entitlements and the attitude of enforcing such entitlements.[41]

The phrase 'freedom of contract' is hackneyed in the writings of contract law scholars, particularly those of apparently libertarian bent. Some of such scholars have taken the view that contracts are outcomes of private orderings reflecting the choices or will of autonomous entities.[42] Others have carried that view even further, holding that contracts should be enforced even though the enforcement of certain contracts may cause untoward outcomes to befall disfavoured parties.[43] Such views only reflect a romantic view of contracting, especially as the bulk of commercial contractual agreements are one-sided or boilerplate, and less likely to be an outcome of thorough negotiations.[44]

These scholars also appear to take for granted the ways of the marketplace as though they were only contingent on the conditions of the 'invisible hand' – i.e. the outcomes of competing entities in the market characterised by different incentives, information asymmetries, varying levels of bargaining power, gradations of opportunity cost and future uncertainties.[45] The assumption is that, with the invisible hand working itself out under the forces of demand and supply, buyers and sellers (i.e. transacting parties) will exchange goods and services, enabling the re-allocation of resources to those who value them most. This view of the marketplace is blinkered. It is so because it does not take account of the system (i.e. the rules of the game) that facilitate the workings of the marketplace and its outcomes, including contracts.[46] The market takes shape against the

40 Ibid.

41 See Warren J Samuels, Marianne F Johnson, and William H Perry, *Erasing the Invisible Hand: Essays on an Elusive and Misused Concept in Economics* (1st edn, Cambridge University Press 2011); see also, Kaushik Basu, *Beyond the Invisible Hand: Groundwork for a New Economics* (1st edn, Princeton University Press 2016)

42 See, for example, Charles Fried, *Contract as Promise: A Theory of Contractual Obligation* (2nd edn, Oxford University Press 2015); see also Ernest J Weinrib, 'Corrective justice in a nutshell' (2002) 52 University of Toronto Law Journal 349; see also, Nathan Oman, *The Dignity of Commerce: Markets and the Moral Foundations of Contract Law* (1st edn, University of Chicago Press 2017).

43 See Paul Davies, 'Bad bargains' (2019) 72 Current Legal Problems 253–286; see also Jonathan Morgan, *Contract Law Minimalism A Formalist Restatement of Commercial Contract Law* (1st edn, Cambridge University Press 2013).

44 See Omri Ben-Shahar (ed), *Boilerplate: The Foundation of Market Contracts* (1st edn, Cambridge University Press 2007); see also, Russell Korobkin, 'Bounded rationality, standard form contracts, and unconscionability' (2003) 70 University of Chicago Law Review 1203; see also W David Slawson, 'Standard form contracts and democratic control of lawmaking power' (1971) 84 Harvard Law Review 529.

45 See Albert Choi and George Triantis, 'The effect of bargaining power on contract design' (2012) 98 Virginia Law 1665; see also Lars Stole, 'The economics of liquidated damage clauses in contractual environments with private information' (1992) 8 Journal of Law, Economics and Organisation 582.

46 See Warren Samuel, 'The economy as a system of power and its legal bases: the legal economics of Robert Lee Hale' (1973) 27 University of Miami Law Review 261; see also E J James and F W Taussig, 'The state as an economic factor' (1886) 7 Science 485.

backdrop of institutional arrangements, at the core of which are legal rules. It is these rules that allocate entitlements (i.e. who owns what). To a large extent, the law's award of entitlements shapes the directions of the invisible hand, in turn, the outcomes in the marketplace. As Robert Hale observed:

> Most of our present distribution of wealth is the result of the relative power, latent or active, of various individuals and groups. The power itself is derived in part from the law's more or less blind and haphazard distribution of favors and burdens, in the shape of powers over others and obligations to others.[47]

At the core of Hale's reasoning is that the bane of distributive inequalities, even in contractual transactions, is the factor of market power, which is the upshot of institutional design or laws. Other eminent legal scholars of heterodox bent such as Richard Ely[48] and John Commons[49] rationalised contract law and transactional outcomes along the same lines as Hale. However, in recent times, Eric Posner and Glen Weyl have sought to inspire a resurgence of similar heterodox reasoning, identifying increased (and undue) market power resulting from private entitlements as the bane of distributive inequalities.[50] They suggest that a radical departure from the 'traditional' conception or 'usual' workings of markets should be pursued to correct this state of affairs. This can be done, they suggest, by designing legal rules with a view to advancing 'free exchange disciplined by competition and open to all comers'.[51]

Libertarians argue against government intervention in the market to disrupt the (supposed) natural workings of the market. However, as Barbara Fried reported Hale to have observed, 'when the government intervened in private market relations to curb the use of certain private bargaining power, it did not inject coercion for the first time into those relations. Rather, it merely changed the relative distribution of coercive power.' [52] Of course, one cannot deny that factors resulting from social interactions such as science, technology, innovation and novel business ideas may shape trends in the marketplace; yet, these factors rely on the forces of legal rules to have a meaningful impact on the market.[53] The law's recognition of rights such as trade secrets, patents, copyrights and other forms of industrial entitlements, exonerates this claim. The law's attitude towards these factors, whether proactive, neutral or negative, is usually informed by the calculations of lawmakers (i.e. legislator or judges) towards social welfare enhancement. In effect, it is institutional arrangements that chiefly determine the marketplace and its outcomes.

The rules of contract law are part of the institutional arrangements that shape outcomes in the marketplace. These rules can have implications for wealth distribution between contracting parties. An excellent example of this is the remoteness rule in the computation of expectation damages in contract law. This rule can be designed to take different forms to place an increased burden on either the promisee or promisor. For example, the principle could take the *Hadley v Baxendale*[54] form in which the promisor is

---

47 Robert Hale, 'Law making by unofficial minorities' (1920) 20 Columbia Law Review 451.

48 Richard Ely, *Property and Contract in their Relations to the Distribution of Wealth* vols 1 and 2 (Macmillan 1914).

49 John Commons, *Legal Foundations of Capitalism* (1st edn, Macmillan Company1924) 65–134.

50 Eric Posner and Glen Weyl, *Radical Markets: Uprooting Capitalism and Democracy for a Just Society* (1st edn, Princeton University Press 2018)

51 Ibid xvii.

52 Barbara Fried, *The Progressive Assault on Laissez Faire* (1st edn, Harvard University Press 2001) 36.

53 See Katharina Pistor, *The Code of Capital: How the Law Creates Wealth and Inequality* (1st edn, Princeton University Press 2019) 11; see also Julie Cohen, *Between Truth and Power: The Legal Constructions of Informational Capitalism* (1st edn, Oxford University Press 2019).

54 (1854) 23 LJ Ex 179.

generally liable for foreseeable losses. The implication of this will be that promisors would have to exclude liability upfront, to avoid legal responsibility.[55] The rule could, yet, take the form propounded by Lord Hoffmann in the case of *Transfield Shipping Inc v Mercator Shipping Inc*,[56] in which mere foresight of the promisee's exposure to loss is not sufficient to allocate risk, rather the promisor's liability must be understood in terms of contractual interpretation and commercial customs. This approach too has its distributive implications. It places a burden on the promisee to secure the promisor's assumption of liability. The upshot of this can vest the promisor with bargaining powers, depending on the transactional context.

Scholars have argued against devising private law rules to achieve distributive effects. At the core of their submission is the position that private law rules are a poor avenue by which one may pursue such objectives; as doing so is apt to distort the workings of the free market.[57] They suggest reliance may be placed on taxation and other fiscal measures instead; as they are less disruptive of the marketplace. Such a position is reminiscent of the parochial *laissez-faire* view of the market, blind to the roles of institutions in commerce. The design of legal rules is always products of considerations that take priority on the minds of lawmakers, the distributive implications of which may or may not be immediately apparent. That is ineluctably the case with private law rules, including those of contract law. But then, what makes private law a poor avenue for achieving distributive ends as claimed? The answer is: nothing! Private law rules can be just as viable as other institutional regimes for advancing distributive goals.[58] The expectation is that lawmakers should duly study the situation they seek to address; select the right response in the form of legal rules; and ensure that the vital incentives of other entities needed towards productivity would not be significantly affected in negative terms.

Contract law rules can be designed or reformed with specific distributive goals in mind, as the illustration above using the remoteness rule shows. In the case of agreed remedies clauses, the distributive role played by judicial intervention is to ensure against the use of contracts as devices for wealth (or rent) extraction; instead of that, they should serve as tools for wealth creation and private empowerment. When contracts can be used by a party to extract ('unearned') wealth, an array of social costs can result.[59] Some may argue that wealth extraction through contracting should not be a problem as it is an outcome of consensual agreements which turned out to be unfavourable to one party. Another argument could be that the enrichment resulting to a contracting party from the ill-advised bargains of another does not dampen social welfare, except that it transfers wealth from one party to another. While these arguments may appear persuasive on the surface, they are undiscerning of the possible use of contracts to squeeze out earnings from an entity without corresponding 'economic sacrifice', creating risks of social costs. Inauspicious states of this kind are particularly possible with agreed damages clauses.

55 See John Barton, 'The economic basis of damages for breach of contract' (1972) 1 Journal of Legal Studies 277; see also, Jason Scott Johnston, 'Strategic bargaining and the economic theory of contract default rules' (1990) 100 Yale Law Journal 615

56 [2008] UKHL 48; [2009] AC 61.

57 See, Louis Kaplow and Steven Shavell, 'Why the legal system is less efficient than the income tax in redistributing' (1994) 23 Journal of Legal Studies 667; see also W N R Lucy, 'Contract as a mechanism of distributive justice' (1989) 9 Oxford Journal of Legal Studies 132.

58 Daphna Lewinsohn-Zamir, 'In defense of redistribution through private law' (2006) 91 Minnesota Law Review 326; see also Anthony T Kronman, 'Contract law and distributive justice' (1980) 89 Yale Law Journal 472.

59 See Barbara Fried, 'Wilt Chamberlain revisited: Nozick's "justice in transfer" and the problem of market-based distribution' (1995) 24 Philosophy and Public Affairs 226.

Understandably, contracting parties may be deft at negotiating bargains and able to secure for themselves price terms that qualify as extortionate. That is an issue which is not the focus of this article. It is better to leave price governance issues to administrative law, particularly the regime of competition law. However, as regards terms of contracts determining measures for enforcing contractual bargains, such provisions fall, undoubtedly, within the purview of contract law's domain. Private entities should not be able to usurp such judicial roles; particularly roles that serve to reflect society's position on the methods and extent of contractual enforcement.[60] The need for legal resistance of such private usurpation of judicial roles is most ripe in our current economic milieu, which, as explained in the introduction, is characterised by vast gulfs in bargaining powers owing to pervasive oligopolistic market structures/market concentration, organised commercial networks, digitisation and big data.

How may we reconceptualise contract law to address this concern? The answer lies in recognising contracts as arrangements for private empowerment. Therefore, to the extent that contractual terms and their execution transcend the lines of private empowerment, they call for judicial scrutiny. Where they extend beyond the lines of private empowerment, they must be justifiable on other defensible grounds parallel to empowerment.[61] The obvious truth is that judges, in establishing common law rules, are not agnostic of the likely distributive implications that legal rules may have. In this regard, in the governance of contracts, the common law has not been shy in the use of default rules or standards as presumptions of the intention of parties concerning how they intend their contracts to be enforced. For example, where an express contract term entitles a promisee to terminate the contract upon a breach by the promisor, the promisee would be disentitled to lost bargain damages for terminating pursuant to such a term, except where the breach relates to a condition or its effect deprives the promisee benefits essential to the contract.[62] Such a rule is part of contract law's institutional control of private ordering.[63] The obvious purpose or implication of such a rule is to discourage the inclination to terminate contractual arrangements, except where it is vital to do so or where the common law right to terminate has arisen. Even where the common law right has arisen, courts are likely to fetter the decision to terminate where it is exercised with a view to gaining an unfair economic advantage or exercised in 'bad faith'.[64] This is in line with the growing acceptance of the precept that contracting parties should be faithful to the purposes or goals of their bargain.[65]

---

60 See Seana Shiffrin, 'Remedial clauses: the overprivatization of private law' (2016) 67 Hastings Law Journal 407, 419–422; see also Ian R Macneil, 'Power of contract and agreed remedies' (1962) 47 Cornell Law Quarterly 495; see also Selene Rowan, *Remedies for Breach of Contract: A Comparative Analysis of the Protection of Performance* (1st edn, Oxford University Press 2012) 230.

61 See, Robin Kar, 'Contract as empowerment' (2016) 83 University of Chicago Law Review 760.

62 See, *Rice v Great Yarmouth Borough Council* [2003] TCLR 1; see also *Bunge Corp v Tradax Export SA* [1981] 1 WLR 711; see also, *Antaios Compania SA v Salen Rederierna* [1985] AC 191; see also, academic articles by John Randall, 'Express termination clauses in contract law' (2014) 73 Cambridge Law Journal 113–141; and also, Richard Hooley, 'Express termination clauses' in Graham Virgo and Sarah Worthington (eds), *Commercial Remedies: Resolving Controversies* (1st edn, Cambridge University Press 2017) 343–365.

63 John Carter and Wayne Courtney, 'Unexpressed intention and contract construction' (2017) 37 Oxford Journal of Legal Studies 326, 382; see also, John Carter and Wayne Courtney, 'Breach of condition and express termination right: a distinction with a difference' (2017) 133 Law Quarterly Review 395.

64 Richard Hooley, 'Controlling contractual discretion' (2013) 72 Cambridge Law Journal 65; see also, Roger Brownsword, 'Bad faith, good reasons and termination of contracts' in J Birds, R Bradgate and C Villiers (eds), *Termination of Contracts* (Wiley Chancery 1990) 283.

65 See *Lymington Marina Ltd v MacNamara & Ors* [2007] NPC 27, paragraph 44; see also, *SNCB Holding v UBS AG* [2012] EWHC 2044 (Comm).

This article advances a case for increased judicial intervention in the enforcement of remedial terms on the rationale that doing so would help, to borrow the words of Posner and Weyl, in enabling 'free exchange disciplined by competition and open to all comers'. This article founds its case for reconceptualising contracts as empowerment arrangements on the premise that contracts are cooperative games.[66] Such theory does not prevent the need for self-interest, nor does it seek to foist altruism on contractual parties. Instead, it advances the argument that contracts are not avenues for opportunistic enrichment.

As postulated by Mark Gergen, contract law should advance two principles, and these are *unselfish performance* and *loss alignment*.[67] By unselfish performance, contract law rules must 'eliminate the incentive a promisor might have to perform inefficiently or that a promisee might have to induce or falsely claim breach if the promisor had to pay damages on breach that greatly exceeded the promisee's loss on breach'.[68] However, by loss alignment, we must tailor compensation to ensure that the promisee is not left worse off than if they had not entered into the bargain in issue.[69] These two principles, although not prominently recognised in the common law of contract, are not alien to the common law of contract and are, in fact, aligned with the incorporation of the ideals of good faith into contract law. This is particularly so as courts are increasingly warming up to the reality that contracts are not adversarial or zero-sum games but relational arrangements for cooperation towards the actualisation of the ends to which bargains relate.[70] As reality shows, the transactional landscape is awash with long-term relations or, at least, repeat contractual relations. The factors of trust and confidence play significant roles in the design, revision and execution of most contracts. Hence, contractual formation or re-negotiation is often 'incomplete' (i.e. not exhaustively providing terms that cover all likely variables and future contingencies).[71] Factors often responsible for contractual incompleteness are information costs, transaction costs and bounded rationality. It is for this reason that scholars often regard incomplete contracts as relational contracts, being informed by evolving mutual understandings that go beyond documentary description.[72]

A characteristic feature of relational contracting is the high switching costs (i.e. the cost of transferring to an alternative) it is likely to impose on contracting parties, with which comes heightened exposure to economic hold-up (or opportunism). Modern judicial attitude mirrors this by incorporating a duty to act in good-faith into contracts – an obligation which is not circumscribed to dealing honestly but extends to being

66 David Campbell, 'The relational constitution of remedy: co-operation as the implicit second principle of remedies for breach of contract' (2005) 11 Texas Wesleyan Law Review 455; see also, Clayton Gillette, 'Commercial relationships and the selection of default rules for remote risks' (1990) Journal of Legal Studies 535.

67 Mark Gergen, 'A defense of judicial reconstruction of contracts' (1995) 71 Indiana Law Journal 45, 46.

68 Ibid 46.

69 Ibid 46–47.

70 George Leggatt, 'Negotiation in good faith: adapting to changing circumstances in contracts and english contract law', Jill Poole Memorial Lecture (Aston University, 19 October 2018) <https://www.judiciary.uk/wp-content/uploads/2018/10/leggatt-jill-poole-memorial-lecture-2018.pdf>; see also Zhong Xing Tan, 'Keeping faith with good faith? The evolving trajectory post-Yam Seng and Bhasin' (2016) 5 Journal of Business Law 420; see also Wayne Courtney, 'Good faith and termination: the English and Australian experience' (2019) 1 Journal of Commonwealth Law 1.

71 Robert E Scott, 'The case for formalism in relational contract' (2000) 94 Northwestern University Law Review 847.

72 Steven Shavell, *Foundations of Economic Analysis of Law* (Harvard University Press 2009) 292–293.

cooperative.[73] It is understandable, however, that not all agreements accommodate good-faith expectations. Such a broad-brush imposition of a duty to act in good faith may be at variance with the essence of certain contractual contexts; typically, gambling, speculation and futures contracts.[74] One way of dealing with contractual incompleteness and uncertainty is the use of agreed remedies clauses to insure against unforeseen eventualities, and indiscriminate deprivation of effect to such clauses could frustrate such commercial engagements.

Notwithstanding the insurance value of such terms, they have potential demerits. They can be used to achieve opportunistic enrichment that has no bearing on loss arising from a contractual breach. They can be used to facilitate market concentration by serving as market entry barriers.[75] They can serve as backdoor specific performance, with the implication that they could bind promisors hand and foot, exposing them to exacting demands. Then, if contract law judges, lawyers and scholars are aware of the relational nature of contracts, why does this reality not feature in how agreed remedies clauses are perceived and enforced?

In the first instance determination of *MSC Mediterranean Shipping Company SA v Claimant Cottonex Anstalt*,[76] Leggatt J (as he then was) introduced the consideration of relational contracting into the governance of remedial provisions. However, in the English Court of Appeal, while agreeing with Leggatt J's decision, his reliance on good faith to reform the remedial provision was rejected. The Court of Appeal's consideration was that Leggatt J could have founded his decision on the basis that there was a frustration of the commercial adventure between the parties to the suit.[77]

The submission of this article is that the treatment of contracts as relational cooperative games should inform the judicial attitude towards intervening in remedial provisions, and that unavoidably requires an incorporation of good faith ideals into contract law. With such an approach, not only would the interests of contracting parties be well balanced, but a proper adjustment to the interests of society would be well-factored into the enforcement of remedial terms.

3 The second strand: the specious categorisation of agreed remedies clauses

Much of the analytical confusion and intricacies surrounding agreed remedies stems from formalistic categorisation. Liquidated damages terms are considered terms that prescribe a sum or a formula for compensation of loss attributable to a specified contractual event.[78] Forfeiture clauses are clauses that entitle a promisee to deprive another or foreclose on an interest owned by another, where that other person has failed to measure

73 See, *Alan Bates and Others v Post Office Ltd* [2019] EWHC 606 (QB); see also, *Yam Seng Pte v International Trade Corp* [2013] EWHC 111.

74 See Mark Gergen, 'A defense of judicial reconstruction of contracts' (1995) 71 Indiana Law Journal 45, 47: 'Unexpectedness is an explicit element of the principle of loss alignment because people do enter into contracts that are zero-sum games, where one party's loss is the other's gain. That is the nature of gambling and insurance.'

75 See Philippe Aghion and Patrick Bolton, 'Contracts as a barrier to entry' (1987) 77 American Economic Review 388; see also Joseph Brodley and Ching-to Albert Ma, 'Contract penalties, monopolizing strategies, and antitrust policy' (1993) 45 Stanford Law Review 1161.

76 [2015] EWHC 283 (Comm).

77 *MSC Mediterranean Shipping Company SA v Cottonex Anstalt* [2016] EWCA Civ 789 (27 July 2016).

78 See *Phonographic Equipment (1958) Ltd v Muslu* [1961] 1 W.L.R. 1379; see also *Edgeworth Capital (Luxembourg) SARL v Aabar Investments PJS* [2018] EWHC 1627 (Comm).

up to contractual expectations.[79] It could be the deprivation of the total value or a portion of a deposit, part payment, or mortgaged interest. Termination clauses entitle a promisee to sever contractual relations with a promisor, thus ending prospective obligations.[80] Termination fees clauses are another, which may additionally be imposed upon effective communication of intention to terminate a contract. There are other recognised labels of ancillary provisions, which cannot all be enumerated here for reasons of space.

The artifice of these categorisations becomes discernible when one focuses on the consequences of these clauses rather than their form. They all share one commonality, and this is that they serve remedial purposes. Their formalistic differences are informed mainly by the factors of legal history and the technicalities of legal practice. But, as we know, life is not rivetted to history and reality betrays formality. Therefore, there should be an abandonment of their classification, and focus should be placed simply on inquiries as to whether such terms are remedial; and, if they are, whether they cross the lines of empowerment and proportionate protection essential to the promisee's interest.

Some scholars of contract law have exposed the inadequacies of such categorisations, particularly concerning the age-long distinction between forfeiture clauses and liquidated damages.[81] Concerning these two forms, they seem to ask: 'What difference is there between an obligation to pay or transfer a sum or asset upon an event and a prescribed sum or asset that the promisor is expected to relinquish to the promisee?' In substance, there is no difference. Courts have recognised the emptiness of such differentiation.[82]

The case of *Nutting v Baldwin*[83] highlights how a clause depriving an errant party of contractual benefits may amount to forfeiture, even where there is neither the deposit of money nor asset as security. The claimants, in this case, sought to contest a decision by the committee of an association to deprive them of the benefits of a successful court suit against certain persons, the purpose for which the said association was formed. The committee of the association had levied additional subscription fees on its members. It also reached a resolution that members who failed to pay the additional fees would be deprived of the benefits of the proposed legal suits. When the association succeeded with the lawsuits planned, it decided to deny the claimants a share in the judgment payment owing to the failure of the claimants to pay the additional levies, which were meant to fund the association's litigation costs. The court recognised that the decision of the committee amounted in substance to a forfeiture of benefits the claimants were otherwise entitled to receive, but the court did not grant relief to the claimant on the reasoning that the committee's decision was legitimate as its purpose was to avoid a derailment of the association's objectives.

As this case shows, the supposed differences between categories of remedial clauses are formalistic, as a contractual termination clause may have the same effect as forfeiture clauses. It may also have the implications of a liquidated damages provision, where a promisor seeking to avoid the deprivation of a contractual benefit or the enforcement of a termination clause might have to pay a sum or suffer a forbearance in order to avoid the promisee making a decision that might have detrimental effects for his or her business

79 *Union Eagle Ltd v Golden Achievement Ltd* [1997] AC 514, 519–523.

80 *Dalkia Utilities Services plc v Celtech International Ltd* [2006] EWHC 63 (Comm).

81 See James Edelman (ed), *McGregor on Damages* (20th edn, Sweet & Maxwell 2018) para 16-014; Hugh Beale (ed), *Chitty on Contract* vol 1 (32nd edn, Sweet & Maxwell 2017) paragraph 26-216N.

82 See *Jobson v Johnson* [1989] 1 WLR 1926.

83 [1995] 1 WLR 201.

interests. In recent times, courts have recognised how private law power derived from contracts may significantly affect the allocation of burdens and benefits between contracting parties. Such powers may be as potent in their effects as those of public authorities.[84] We can imagine the suspension of a bank from a payment card network; the shutting-out of a company from a consortium blockchain network which facilitates its business activities; or the termination of a supplies contract where a bulk buyer has minimal alternatives sources of supply to fall upon, to mention a few.[85] In similar vein, a body of case law has applied administrative law reasoning to assess the reasonableness or otherwise of discretionary decisions of parties wielding private law powers.[86] This is pertinently so in those cases where contractual terms entitle promisees to terminate contracts for breach, and the courts have interpreted such powers to be akin to discretionary powers that should be exercised in good faith.[87] But if there is an appreciation of this reality in case law, why has it not changed judicial rules concerning relief from forfeiture? We shall return to this discussion below when addressing relief from forfeiture clauses and the bias against purely contractual interests.

The unappreciated implication of practical legal decisions, like those of the Australian High Court in *Andrews* and *Pacciaco*, is the doing-away with these needless categorisations, followed by a uniform treatment of remedial terms. Halson reasons that an attempt at drawing a commonality between agreed remedies clauses, particularly stipulated sum clauses and those of forfeiture of deposits, is simplistic.[88] He posits that this is so for two reasons. One is the issue of timing. An assessment of the validity of a stipulated sum clause as being against the parties' expectations is viewed at the time of contracting; while that of forfeiture clauses is at the time of enforcement. The second relates to considerations that inform the said assessment. That the party seeking to enforce stipulated sum provisions must show that the term serves as 'proportionate protection to a legitimate interest of the promisee';[89] while those seeking to enforce a forfeiture clause only need to show that retention of the promisor's interest is reasonable. This second basis of distinction is devoid of substance. We shall return to this below when addressing the issue of legitimate interest. However, the first basis of differentiation raised by Halson (that of timing), while right on account of legal doctrine, is also destitute of substance. Thus, where there is an expurgation of the timing difference, the divergence between both clauses evaporates. Below, we shall return to the discussion on the imperative to discard the timing difference.

84 See Hon Justice Stephen Kós, 'Constraints on the exercise of contractual powers' (2011) 42 Victoria University Wellington Law Review 17; see also Robert Hale, 'Coercion and distribution in a supposedly non-coercive state' (1923) 38 Political Science Quarterly 470.

85 See, generally, Jose Gomez-Ibanez, 'Regulating infrastructure: monopoly, contracts and discretion' (1st edn, Harvard University Press 2003).

86 See Michael Bridge, 'The exercise of contractual discretion' (2019) 135 Law Quarterly Review 227; see also *Braganza v BP Shipping Ltd & Another* [2015] UKSC 17; *British Telecommunications plc v Telefónica O2 UK Ltd* [2014] UKSC 42, [2014] Bus LR 765; *Socimer International Bank Ltd v Standard Bank London Ltd* [2008] EWCA Civ 116; *Paragon Finance plc v Nash* [2001] EWCA Civ 1466, [2002] 1 WLR 685; *Abu Dhabi National Tanker Co v Product Star Shipping Ltd (The 'Product Star') (No 2)* [1993] 1 Lloyd's Rep 397.

87 *Lymington Marina Ltd v MacNamara & Others* [2007] NPC 27, paragraph 44; see also, *SNCB Holding v UBS AG* [2012] EWHC 2044 (Comm).

88 Roger Halson, *Liquidated Damages and Penalty Clauses* (1st edn, Oxford University Press 2018)156–157.

89 Ibid.

4 The third strand: considerations for recalibration

There are numerous legal variables associated with remedial terms. However, some are predominantly salient. It is these notable elements of remedial terms that this section addresses in advancing the case for recalibration. This article separates these elements into different heads for analysis. The heads addressed here are as follows:

1. the breach requirement;
2. the legitimate interest question;
3. absence or otherwise of a duty to mitigate;
4. whether a promisor has the right to elect between the enforcement of remedial clauses and falling back to common law damages
5. *ex-post* considerations; and
6. relief against forfeiture and the bias against contractual rights.

4.1 The pointlessness of the breach requirement

In modern contract law, the predominant position appears to be that a breach is a precondition for the rule against penalty to apply. The decisions of the UK Supreme Court in *Cavendish* and the High Court of Australia in *Andrews* (reaffirmed in *Pacciaco*) revived this question. The former court held that breach is a precondition, while the latter ruled that breach is not a precondition. There have been varying academic explorations of the gulf between these two significant courts, and, as such, this article does not engage in parsing these judgments. That primarily is because this article pursues a prescriptive case against the said breach requirement.

The said requirement rests on the rationale that the jurisdiction to intervene in a remedial clause only arises where the obligation to pay is one of a secondary nature coming into effect upon the breach of a primary one. Allied to this, however, are the rules that judicial intervention would not arise in cases where a secondary obligation to pay is not for the benefit of the party to whom the promisor owes a primary responsibility, but for a third party;[90] nor does it apply upon a specified event that has no bearing on breach[91] – e.g. where an identified entity goes bankrupt, or any external occasion arises.

It is not merely that the breach requirement is problematic because distinguishing between primary and secondary obligations can be an impossible task in the face of innovative and creative contract drafting.[92] It is that it is inherently inadequate in allaying concerns around the use of contracts to achieve an undeserved, upward distribution of wealth and wealth extraction as enabled by substantial degrees of bargaining power. What courts should be concerned about, when dealing with remedial clauses, is what the essence of the bargain between the parties to a contract is. As some judges have recognised, in determining whether a provision is of remedial value, an undue focus should not be placed on textual interpretation of contractual documents but on the

90 See *Export Credits Guarantee Department v Universal Oil Products Co* [1983] 1 WLR 399; see also *Imam-Sadeque v BlueBay Asset Management (Services) Ltd* [2012] EWHC 3511 (QB).

91 See *Euro London Appointments Ltd v Claessens International Ltd Edgeworth Capital (Luxembourg) SARL and Another v Ramblas Investments BV* [2015] EWHC 150 (Comm).

92 See David Foxton, 'How useful is Lord Diplock's distinction between primary and secondary obligations in contract?' (2019) 135 Law Quarterly Review 249; see also Justice Julie Ward, 'Penalties and the protection of freedom of contract' (Banking and Financial Services Law Association Conference, 15 August 2020) <http://www.austlii.edu.au/au/journals/NSWJSchol/2010/18.pdf>.

matrix of fact peculiar to the parties at the time of forming the contract. For as Lord Radcliffe expressed it in *Bridge v Campbell*:[93]

> The court's jurisdiction to relieve against penalties depends on 'a question not of words or forms of speech but of substance and of things' … It cannot really depend on a point of construction, though it is often spoken of as so depending. A sum of money sued for in one set of circumstances, as on a hirer's breach, when alone the 'in terrorem' idea can have any application, may be a penalty in the eyes of the law, without it being necessarily anything but the price of an option in another set of circumstances or a mere guarantee in yet a third.[94]

Elaborate schemes may be adopted to defeat the formulaic breach requirement.[95] For example, suppose B (a promisor) intends securing a contract with A (a promisee) in circumstances where A is sceptical about B's reliability. A requires B to obtain a letter of credit from a bank in favour of himself (A) or any of his nominees, with the instruction to the bank that where A or any of his said nominees can show that a particular task imputed to B has not been performed or if a given event occurs, then the bank should pay A or the said nominees $X. Should B not perform, or the said event happens, and the bank pays the elected payee, B's chances of contesting the payment as a penalty would be bleak if we were to go by the breach requirement for the following reasons. One is that B's standby instruction to the bank preceded B's default. The second is that the bank's payment to A or his nominee was not made by B to A but through a third party (i.e. the bank). Thirdly, where the payment is to A's nominee with whom B has no contractual relations, B's chances are most dim; as there would be no breach of a contractual obligation owed to the payee.

The analogy is equally applicable with the same implications in cases where A and B had adopted an escrow arrangement, with A paying in advance to a third party who is instructed to pay B should there be breach or a specified event. Thus, it is patently clear that only by an understanding of the real bargain between parties that we can assess the true purpose of the said payment arrangement to determine if it serves a remedial role. In a similar vein, the Scottish Law Commission in a recently issued report on remedial clauses described the breach requirement as being too narrow on the ground that 'it would not catch the payment to be made when a party exercised an option under a contract such as terminating it early'.[96]

There are fears expressed concerning the call for dispensing with the breach requirement. A common one is that it carries with it the risk of disrupting contract structuring, heightening difficulties for judges and lawyers. Another concern is that discarding the breach requirement creates room for commercial uncertainty, along with the effect of enabling contractually disadvantaged parties to be freed from ill-advised bargains.[97] Such views appear appealing on the surface; however, they only perpetuate the case for form over substance. We can allay such fears by stating the requirement for the enforceability of remedial clauses as the demonstration of a legitimate commercial interest for protection. The imposition of such a condition does not create commercial uncertainty;

---

93 [1962] AC 600.

94 Ibid 624; see also *Lombank Ltd v Excell* [1963] 3 WLR 700.

95 See Gerald McLaughlin, 'Standby letters of credit and penalty clauses: an unexpected synergy' (1982) 43 Ohio State Law Journal 1.

96 Scottish Law Commission, *Review of Contract Law: Discussion Paper on Penalty Clauses* (Discussion Paper No 162, 2016) 10.

97 See John Carter, Wayne Courtney, Elisabeth Peden, Andrew Stewart and G J Tolhurst, 'Contractual penalties: resurrecting the equitable jurisdiction' (2013) 30 Journal of Contract Law 99, 112–113.

it only serves to adjust the balance of legal entitlements between contracting parties. We shall return to this issue when addressing 'legitimate commercial interest' below.

What should be the criterion for identifying a clause as being of remedial implications? We can answer that question by referring to a statement of Deane J in his dissenting judgment in the Australian case of *AMEV-UDC Finance Ltd v Austin*,[98] where he said, concerning the identification of remedial clauses, that:

> The general area in which they are applicable is where there exists a contractual liability … to pay or forfeit an amount or amounts either on or in default of the occurrence of an event which can be seen, as a matter of substance, to have been treated by the parties as lying within the area of obligation of the party liable to make the payment in the sense that it is his or her responsibility to ensure that the specified event does or does not occur and where the stipulated payment contains an element of compensation for the economic loss or damage which might be sustained by the other party by reason of the particular occurrence or default.[99]

Thus, the baseline of the inquiry should be whether there is an identifiable responsibility or event which the bargain of the parties requires the promisor to facilitate or avoid. Then a collateral obligation upon the promisor (whether to pay or do something) arising in response to the unwanted outcome would be of remedial implication where it serves to rectify the said undesirable outcome. Ultimately, however, whether the stipulated measure in response to the undesired outcome is remedial can only be determined by a scrutiny of the matrix of facts between the contracting parties at the time of contractual formation. Such exercise would help us ascertain what interests were at stake in the parties' agreement. By so doing, we can distinguish substantive terms (e.g. options and prices) from collateral ones that serve to remedy unwanted events. Without such a careful approach, we are bound to fall into error. For, as admonished by Lord Halsbury in *Castaneda and Others v Clydebank Engineering and Shipbuilding Co Ltd*:[100] 'it is impossible to lay down any abstract rule as to what it may or it may not be extravagant or unconscionable to insist upon without reference to the particular facts and circumstances which are established in the individual case'.[101]

4.2 The legitimate interest condition

Ever since the decisions of *Cavendish*, *Andrews* and *Pacciaco*, the phrase 'legitimate interest' has gained prominence as the new touchstone for determining the validity of agreed remedies clauses across commonwealth jurisdictions.[102] Thus, the new standard is whether the object of the clause is out of proportion with any legitimate interest which the promisee sought to protect.[103] Although the jurisprudence on the factor of legitimate interest concerning remedial clauses is still nascent, we can draw guidance on what counts as a legitimate interest from an analogous aspect of contract law in which the condition

98 (1986) 162 CLR 170

99 Ibid paragraph 5 of Deane J's dissenting judgment.

100 (1904) 12 SLT 498.

101 Ibid 500.

102 See, for example, the New Zealand Court of Appeal decision in *Honey Bees Preschool Ltd v 127 Hobson Street Ltd* [2019] NZCA 122; see also the Singaporean High Court decision in *Seraya Energy Pte Ltd v Denka Advantech Pte Ltd* [2019] SGHC 2.

103 See Lord David Hope, 'Law of penalties – a wasted opportunity? (2016) 33 Journal of Contract Law 93.

serves as a crucial requirement.[104] What one may glean from discussions bordering on legitimate interests is that they are interests vital to protecting the commercial and economic objectives of the promisee following the essence of the bargain between the parties. As one scholar describes it, it must have a nexus with the promisee's performance interest.[105] It follows that where a promisee seeks to use the facility of a contract to enrich themselves without any bearing on the essence of the bargain between the parties, then such cannot amount to a legitimate interest.

There appears the position that being able to show a legitimate interest saves a remedial clause from invalidation.[106] That view may be logically valid, but it only gains validity because of assessing enforceability of a remedial provision at the time of contracting, and not at the time of the breach. The submission of this article is that being able to show a legitimate interest should not end the matter concerning enforceability. A term which might appear at the time of contracting to be legitimate may turn out disproportionate in effect at the time of breach or the eventuation of the undesirable event. For example, where the loss to the promisee is significantly lower in value than had been feared by the promisee at the time of contracting. It follows that having a legitimate interest in the use of a remedial clause should not exclude judicial review. Courts should be able to scale down a provision whose effects would be disproportionate to the legitimate interests of the promisee in the light of facts available at the time of breach or an undesired event. This issue is well addressed below under the subheading '*Ex-post* considerations'.

4.3 The duty to mitigate

It appears a settled rule concerning agreed damages clauses that they create a debt obligation in favour of entitled promisees. For this reason, the position is that such promisees have no bilateral duty to take measures towards the mitigation of their exposure to losses. Such an obligation to mitigate only arises, it is argued, in cases of unliquidated damages. The rationale for this rule was in modern times stated in the case of *Abrahams v Performing Rights Society Ltd*[107] and was recently re-echoed by Leggatt J in *MSC Mediterranean Shipping Co*, saying, among other things, that:

> to allow mitigation arguments where there is a liquidated damage clause would be inconsistent and unfair because it would involve limiting the damages recoverable by a plaintiff who can show that his actual loss is greater than the stipulated sum whilst permitting a defendant who can show that it is less to take advantage of that fact. It would also expose the parties to 'the risk, expense and uncertainty of litigation the avoidance of which is to be presumed to be one of the principal reasons for their stipulating for liquidated damages'.[108]

104 Solene Rowan, 'The "legitimate interest in performance" in the law on penalties' (2019) 78 Cambridge Law Journal 148.

105 Jessica Palmer, 'Implications of the new rule against penalties' (2016) 47 Victoria University of Wellington Law Review 287.

106 Elizabeth Peden, 'Penalties after Paciocco – the enigma of "legitimate interests"?' (2019 Journal of Contract Law 263, 277; see also, Carmin Conte, 'The penalty rule revisited' (2016) 132 Law Quarterly Review 382 at 387: 'Once the court identifies a legitimate interest, the difficulty in conducting the comparative exercise in a principled fashion will mean that, in practice, it will always uphold the agreed remedy clause. The rule will have been abolished by the back door.'

107 [1995] ICR 1028

108 [2015] EWHC 283, paragraph 70.

One may try to rationalise the English Court of Appeal's decision in *Murray v Leisureplay*[109] on the absence of a duty to mitigate in liquidated damages cases. However, a proper understanding of that case shows that the court did not identify such a rule. Instead, the court dismissed the argument in favour of a duty to mitigate as it relates to the peculiarities of the case. The basis for that decision was that the promisee–employee (Murray) had a legitimate interest in receiving the contractually stipulated sum even though he had the opportunity to mitigate his exposure to loss of income by seeking employment elsewhere.[110] Most importantly, the essence of the contract in issue between the parties did not give room for the recognition of such a duty to mitigate loss.

That there should be no imputation of a correlative duty of mitigation on a promisee entitled to stipulated damages payments cannot stand as a defensible rule in a milieu in which proof of an absence of legitimate interest may disentitle a promisee from receiving a contract price, as established in the case of *White and Carter (Councils) Ltd v McGregor*.[111] Suppose X is entitled, based on an executory agreement, to receive a price payment from Y on the expectation of a counter-performance by X. X may yet not be entitled to receive such amount as agreed where X is aware of Y's repudiation or lack of capability to perform but continues to keep the contract open. The rationale for rejecting X's entitlement to payment is the absence of any genuine basis for insisting upon contractual performance beyond self-enrichment that cannot fairly be reconciled with the goal of their bargain. In other words, X should not be allowed to use the agreement as an avenue for enrichment. Why then should such 'legitimate interest' requirement not apply to a promisee entitled to a liquidated sum upon breach or similar events, especially in connection with that promisee's ability to mitigate against loss? Except where a promisee can show that the context and essence of the bargain between the parties entitled the promisee to an undiscounted payment of the stipulated sum, the law should require mitigation wherever doing so is inexpensive, and there are low-risk avenues available to the promisee in achieving such mitigation. It does not mean that promisees should assume sacrificial roles, the implication of which would place onerous financial obligations on them.

It may be argued, as foreshadowed in the introduction to this article, that liquidated damages may legitimately serve to predetermine the measure of loss that may arise upon non-performance, especially in situations where losses may be difficult to quantify. It does not, however, follow that it therefore obviates the need for mitigation. This is so because, since the essence of mitigation is to uphold the policy of making compensation approximate the net loss resulting from breach or analogous events, it should applied in reining in the enforcement of liquidated damages. Mitigation does not serve merely to prevent the promisee from accruing additional losses and transferring them to the promisor; it also serves to ensure that a promisee, in line with good-faith expectations, does not frustrate the goals of a contractual bargain.

Two significant reasons justify the recognition of such a duty to mitigate. One is that it avoids the waste of resources, thus promoting economic efficiency in circumstances where the promisee is the better cost avoider.[112] There is a significant risk that a promisee entitled to a liquidated sum may fail to take advantage of inexpensive and low-risk routes

109 [2005] EWCA Civ 963.

110 See *Cavendish Square Holding BV v Makdessi ParkingEye Ltd v Beavis (Consumers' Association Intervening)* [2016] AC 1172, 1269–1270.

111 [1961] UKHL 5, [1962] AC 413.

112 See, Melvin Eisenberg, 'The duty to rescue in contract law' (2002) 71 Fordham Law Review 647.

to reduce economic loss if a duty to mitigate is not recognised.[113] This is most likely so in cases where the stipulated sum is significantly more than the promisee's net loss. However, in cases where the stipulated sum is likely to be lower than the promisee's net loss, the promisee may feel compelled to mitigate losses anyway.

Another reason is that it may enable opportunism on the part of the promisee. The fact that a compensatory sum is fixed does not eliminate the possibility of the promisee acting opportunistically. Thus, in cases where a promisee can avoid loss by taking proactive measures, the non-recognition of a duty to mitigate will vest bargaining power in the promisee where the promisor seeks to implore the former to make efforts towards mitigation. And 'double-dipping' will be implicit in cases where the promisee sues to claim the sum stipulated after they have mitigated loss.[114] Therefore, to avoid placing the promisor at the mercy of the promisee in this regard, it is essential to entrench the legal recognition of contracts as relational cooperative arrangements. The effect of this would be to impose a duty to mitigate loss in cases of liquidated damages, where such is possible at low-cost to the promisee.

One may extend the case for the duty to mitigate to promisees seeking to enforce forfeiture clauses as well. On the authority of *Cukurova Finance International v Alfa Telecom*,[115] a party seeking relief from a forfeiture clause may not rest their case for relief on the basis that the promisee acted in bad faith. That is so even where, for example, the promisee's actions create the risk of jeopardising the efforts of the promisor in avoiding forfeiture.[116] Such a rule is difficult to defend. It is also difficult to reconcile with some of the well-established equitable factors that allow promisors to claim relief. Potentially, the jurisdiction to relieve a promisor from forfeiture is available in (limited) cases, mainly where the objective of the remedial provision in the contract was to facilitate the security of payment to the promisee.[117] However, the promisor must show that, except for the non-fulfilment of the contractual obligation warranting the enforcement of the said clause, they conducted themselves properly; and that they are still able and willing to meet contractual expectations.[118] Also, relief should also be available where there is a significant gap between the value of the interest to be lost in the enforcement of the forfeiture clause and the possible detriment likely to be suffered by the promisee.[119] The conduct of the promisee also counts. For example, relief may be available where the 'conduct of the person seeking to secure the forfeiture has been wholly unreasonable and of a rapacious and unconscionable character'.[120]

A party seeking to enforce such a clause would undoubtedly qualify as acting unreasonably, and certainly unconscionably, where they strive to frustrate the efforts of the promisor in redeeming their interest or actualising the fulfilment of contractual

113 See *Financings Ltd v Baldock* [1963] 2 WLR 359, 363–367; see also *Campbell Discount Co Ltd v Bridge* [1962] 2 WLR 439, [1962] 1 All ER 385, [1962] AC 600; see also *United Dominions Trust Ltd v Ennis* [1968] 1 QB 54. In these cases, one can glean that a duty to mitigate was implicitly recognised. The rationale being that where a promisee had terminated a contract upon the promisor's breach, losses resulting to the promisee from the decision to terminate are attributable to promisee (and not the breaching promisor).

114 See Lisa A Fortin, 'Why there should be a duty to mitigate liquidated damages clauses' (2009) 38 Hofstra Law Review 285.

115 [2013] UKPC 2

116 Ibid 16–17.

117 *Shiloh Spinners Ltd v Harding* [1973] AC 691; see also *Warnborough Ltd v Garmite Ltd* [2003] Civ 1544.

118 See *Gill v Lewis* [1956] 2 QB 1, 13–14; see also, *Inntrepreneur Pub Co (CPC) Ltd v Langton* [2000] 1 EGLR 34.

119 *Shiloh Spinners Ltd v Harding* [1973] AC 691, 725.

120 Ibid 726.

obligations.[121] Such unreasonable actions may take shape in making it impossible for the promisor to obtain finance; failing to disclose vital information, or acting manifestly uncooperatively.[122] However, we must still have regard to the nature of the transactional context between the parties.

4.4 Relief from forfeiture and the bias against contractual rights

The rule appears to be that property and personal property interests are the only recognised candidates for relief against forfeiture.[123] Traditionally, relief against forfeiture was considered exclusively available to interests in real property, such as leases and freehold. That was on the theory that the promisee seeking to enforce forfeiture concerning such rights would be entitled to specific performance and other corrective remedies should the promisor's action trigger the call for such a clause.[124] Then, with the advancement of time, the courts expanded the availability of the relief to interests in personal property, on the reasoning that the distinction between land and personal property is unsustainable.[125] However, as regard interests of a purely contractual nature, such as a licence in intellectual property, or any contractual right not amounting to property or personal property, relief is generally unavailable.[126] Cases in which promisors were able to secure such relief in connection to purely contractual rights have now been rationalised as dealing with interests analogous to personal property.[127]

One may then ask, 'why should relief from forfeiture not be available to promisors with purely contractual interests'? The manifest bias against contractual interests in this regard is unsustainable in the prevailing knowledge-/information-/network-based economic dispensation where contractual interests are just as valuable as property interests.[128] The exercise of the power to terminate contracts or withhold a contractual benefit may significantly alter the allocation of benefits and burdens between parties to an agreement. The most relatable examples are cases of expulsion from business networks, associations or consortiums, and situations of asset-specificity – i.e. where a party has become so invested in a resource that their cost of finding an alternative can only come at an inordinate expense.[129] Relief should also be available where there is the risk of the promisee gaining an economic windfall from insisting on forfeiture in circumstances where cheaper alternative corrective measures are enough.[130] Thus, a

121 See, the statements of Justice Cardozo in the USA case of *Jacob & Youngs, Inc v Kent* 230 NY 239 (1921), 243–245.

122 See, for example, the explanations of Leggatt J in *Yam Seng Pte Ltd v International Trade Corp Ltd* [2013] EWHC 111 (QB), paragraph 142; see also Hugh Collins, 'Implied duty to give information during performance of contracts' (1992) 55 Modern Law Review 556.

123 See *Scandinavian Trading Tanker Co AB v Flota Petrolera Ecuatoriana* [1983] 2 AC 694.

124 See *General Motors UK Ltd v The Manchester Ship Canal Co Ltd* [2018] EWCA Civ 1100, paragraphs 40–47.

125 See *BICC plc v Burndy Corp* [1985] Ch 232, 252A–C.

126 See *SCI (Sales Curve Interactive) Ltd v Titus Sarl; Sport International Bussum BV v Inter-Footwear Ltd* [1984] 1 WLR 776; *Union Eagle Ltd v Golden Achievement Ltd* [1997] AC 514, 519F–G.

127 See *General Motors UK Ltd v. The Manchester Ship Canal Company Ltd* [2018] EWCA Civ 1100; where *On Demand Information plc v Michael Gerson (Finance) plc* [2001] 1 WLR 155 was rationalised as a case relating to personal property. Also, the Court of Appeal rejected the first instance decision on the matter before it at [2016] EWHC 2960 (Ch), which recognised the possibility of relief to contractual interests. Instead, the court rationalised the contractual licence in issue as relating to a personal interest.

128 See Kit Barker, 'Economic loss and the duty of care: a study in the exercise in the of legal justification' in Charles Rickett (ed), *Justifying Private Law Remedies* (Hart 2008) 175.

129 See Chapin Cimino, 'The relational economics of commercial contract' (2015) 3 Texas A&M Law Review 91, 109

130 See *On Demand Information plc and Another v Michael Gerson (Finance) plc* (2001] 1 WLR 155, 172.

promisee should only be able to pursue forfeiture if they can show legitimate interests for enforcing it.

The bias against contractual interests in this regard cannot stand for much longer as courts now demonstrate the willingness to reverse or revise terminations that have punitive implications. An excellent example of this is the case of *Vivienne Westwood Ltd v Conduit Street Development*,[131] where a landlord terminated a contractual concession which allowed the payment of a lower sum as rent, replacing it with a substitute term requiring a higher amount. The landlord did so because the tenant had technically breached the contract by paying late on a given occasion. The court found such termination to be punitive in effect in the circumstances of the case. The court reasoned, most importantly, that the financial implications of the said termination imposed disproportionate financial obligations on the tenant, especially in a situation where whatever injury likely to result to the landlord was small, and less burdensome measures to the promisor could address the promisee's injury.

One should not take this as the calling for judicial rescue from bad bargains. Contracting parties with powers to terminate or revise contracts should be able to do so in the advancement of self-interest in cases where legal limitations would render a bargain commercially obtuse. Such would be the case in situations where judicial intervention could expose the promisee to the risk of economic burdens like the loss of better deals or cost increase.[132] Also, there should be no need for legal intervention where, for example, a higher sum is the real bargain between the parties, but there is concession allowing a lower amount; and the promisee may resume the higher amount upon breach or an undesired event.[133] In such cases, the high amount is the original price of the contract and, thus, not a substantive term which falls under the jurisdiction to review remedial clauses. Additionally, relief should not generally be available in cases of forfeiture of instalments already paid towards the purchase of a subject matter, where such instalments count as part of the contract price, except where the buyer can demonstrate manifest unfairness.[134]

Judicial review of termination clauses and clauses with similar results should be available in cases where the promisee fails in showing a legitimate interest for insistence on such provisions, and there is a real likelihood of incommoding the promisor with heightened burdens. Such would be the case in situations where the promisee's stake does not go beyond money payments; where the object of the contract between the parties is still attainable; there is no situation of irreconcilable break down in commercial relations; and, if there is, the bargain does not require a personal relationship to execute.

4.5 *EX-POST* CONSIDERATIONS

An important rule concerning stipulated sums or assets, which in prevailing law sets it apart from those of forfeiture clauses, is that the time at which the promisee's legitimate interest in claiming the said sum is the time of contracting. Suppose a promisee had a specific commercial interest concerning which a stipulated amount or asset was agreed as insurance against default. Suppose, also, that the promisor's action did offend the said concern, but

131 [2017] EWHC 350 (Ch).

132 See *Socimer International Bank Ltd* [2008] EWCA Civ 116, 122–123.

133 *Wright & Another (Liquidators of SHB Realisations Ltd) v The Prudential Assurance Company Ltd* [2018] EWHC 402 (Ch); *Thompson v Hudson* (1869) LR 4 HL 1.

134 See *Stockloser v Johnson* [1954] 2 WLR 439; *Workers Trust & Merchant Bank Ltd v Dojap Investments Ltd* [1993] AC 573.

the magnitude and value of injury resulting are substantially insignificant as compared to the quantum of both what the promisee had feared and the stipulated sum or the asset the promisor had avowed to pay. The promisor would not by dint of the penalty rule be relieved from the obligation to pay the specified measure notwithstanding the gap in value at the time of the breach. Going by the prevailing rule, so long as the promisee had a discernible commercial fear and a proportionate legitimate interest in protecting that concern at the time of contracting, the said gap in value would not affect the promisee's entitlement to receive the specified sum or asset. As already stated above, there is no difference in substance between this rule and those governing forfeiture clauses; especially as the presentation or drafting of stipulated sums or assets clauses may take creative forms and one form of the clause may be judicially understood as being another.

There is no justifiable basis upon which one may defend such a ('time of contracting') rule. The common defence often presented is that, so long as the parties to the bargain acted freely, and are, possibly, sophisticated entities who are well-advised, the enforceability of such clauses should seldom be subject to review.[135] The weakness of such an argument is discernible. If a fear existed at the time of contracting but dissipates at the time of the breach, or the undesirable event, then it does not make any sense to allow the promisee to keep the windfall accruing on account of that gap. There can be no justification for allowing such a windfall. Such an approach relegates the gist of the contract and treats an appendage bargain (i.e. the remedial clause) as though it were the main thing the parties had in mind. There is an imperative to adjust the appendage bargain to the needs of the contract at the time of the breach. A windfall may, however, be defensible where it serves to protect the legitimate interest of the promisee and there is justification for it based on the commercial context and purpose of the bargain between the parties.

As Eisenberg explains, a piece of reality not present in the conception of the 'time of contracting' rule is that humans have defective telescopic abilities.[136] Contracting parties customarily assume that their current capabilities and calculations will sufficiently avail them towards the performance of contractual obligations. A promisor may be aware of liquidated damages clauses, but the full implications of such provisions are less likely to have been given thorough consideration by him or her. There is the likelihood that where a promisee is confident that the value of a stipulated sum will be more than that of the injury likely to result from a breach, the promisee might have an incentive to encourage such breach.[137] Thus, the promisee might feel emboldened to act uncooperatively towards the promisor, which may equally cause the promisor to spend wastefully towards avoiding a breach. It would also exceedingly tilt the balance of relations between the parties towards the promisee's favour. The case would undoubtedly be different if the incentive to act uncooperatively were removed, which could be done by the judicial adjustment of the stipulated sum towards approximation with the losses ensuing at the time of breach.

135 See, *General Trading Co (Holdings) Ltd v Richmond Corp Ltd* [2008] EWHC 1479 (Comm), paragraph 133; *Philips v Attorney General of Hong Kong* [1993] 61 BLR 41.

136 Melvin Eisenberg, 'The limits of cognition and the limits of contract' (1995) 47 Stanford Law Review 211, 222–228.

137 See Kenneth Clarkson, Roger Miller, and Timothy Muris, 'Liquidated damages v penalties: sense or nonsense' (1978) Wisconsin Law Review 351; see also, Timothy Muris, 'Opportunistic behavior and the law of contracts' (1981) 65 Minnesota Law Review 521.

This risk may be illustrated using a hypothetical example.[138] Suppose, a remedial clause in a joint venture contract for a software development project entitles non-defaulting parties to buy out the interests of defaulting parties at a flat discounted rate of 20 per cent. Before the launching of the software in the market, a defaulting party's interest in the arrangement is worth, say, $100,000. Suppose, however, that within the first year of launching of the software in the market, the value of the defaulting party's share is $600,000. If the promisee seeks to enforce the remedial clause when the value of the promisor's asset was $100,000, the discount value would be $20,000. Such a discount rate is one which the promisee may not have much difficulty in establishing a legitimate interest over. However, were the enforcement of the clause to be upon market entry of the software, when the promisor's share is $600,000 in value, a discount rate of 20% would be $120,000. Thus, it becomes clear that by not adjusting agreed remedial clauses in correspondence with hindsight information regarding the loss of the promisee, undue windfall opportunities may accrue to promisees. The call for adjustment is not that where hindsight shows that the promisee's loss is significantly less than the value of an agreed remedies clause, the court should declare it unenforceable and replace it with common law damages. Instead, the submission of this article is that the value of such provisions can be scaled down to a measure that approximates to the legitimate interest of the promisee, judging by events at the time of the breach. We address this in the next subheading.

4.6 The election between remedial clauses and standard contract remedies at the time of enforcement

There is the rule that where a promisee has stipulated a sum as damages, and the said sum turns out to be lower than the actual injury suffered, that party cannot then elect to fall back on standard contract law remedies. There is also the rule that where stipulated sums turn out to be punitive, they are not just unenforceable but void *ab initio*[139] and may not be judicially revised.[140] The implications of both rules are that promisees are restricted to either common law damages or liquidated damages; and where the latter is chosen, a scaling-down of the stipulated sum is not to be allowed where it qualifies as a penalty. This was succinctly criticised by Treitel, saying:

> The common law rules for distinguishing between penalties and liquidated damages manage to get the worst of both worlds. They achieve neither the certainty of the principle of literal enforcement, since there is always some doubt as to the category into which the clause will fall, nor the flexibility of the principle of enforcement subject to reduction, since there is no judicial power of reduction. On the other hand, they place an undue premium on draftsmanship … the chief danger is to 'home made' clauses which may be invalidated even though they are not intrinsically unfair.[141]

There should be a reformation of these rules based on the theory that agreed damages are in substance the same regardless of categorisation. If a promisee seeking to enforce a forfeiture clause realises that the value of the promisor's security is below that of the

---

138 See Michael Lishman, 'Penalties and relief against forfeiture of joint venture interests' (2008) 27 Australian Resources and Energy Law Journal 219, 234.

139 See *Cavendish Square Holding BV v Makdessi; ParkingEye Ltd v Beavis* (Consumers' Association Intervening) [2015] UKSC 67, [2016] AC 1172, 1220–1222; see also *Wall v Rederiaktiebolaget Luggude* [1915] 3 KB 66, 73; see also *Financings Ltd v Baldock* [1963] 2 QB 104, 120.

140 *Dalkia Utilities Services plc v Celtech International Ltd* [2006] EWHC 63 (Comm), paragraph 128.

141 Guenter Treitel, *Remedies for Breach of Contract: A Comparative Account* (1st edn, Clarendon Press 1988) 233.

likely injury, the promisee may insist on the forfeiture and request for additional sums to make up for the balance of loss. Also, a forfeiture claim may be scaled down to a measure equivalent to the value of the interest lost by the promisee. Why does the law not apply similar approaches to liquidated damages as well?

The first rule above (i.e. that where the value of actual injury turns out higher than that of the sums stipulated, the promisee cannot revert to common law damages) appears to have overtaken the competing position which was that a promisee could return to common law claims where stipulated sums turned out lower than the value of a loss.[142] The reason for the change in rule appears attributable to the rationale of contractual certainty and finality.[143] It is only proper that we perceive liquidated damages as foreshadowing the promisee's exposure to loss, especially as prediction may be challenging to ascertain beforehand and the promisee might have idiosyncratic expectations. In any case, the law should recognise a promisee's entitlement to choose between a stipulated sum and common law damages (as a fallback position) if the former turns out lower than the latter. Also, if a contract does not expressly stipulate an exclusion of unliquidated damages, the law should recognise them as an automatically available fallback for promisees.

The second rule above (i.e. that where a stipulated sum qualifies as a penalty it shall be void *ab initio*) is equally a product of modern innovation as the first. In the relatively recent case of *Jobson v Johnson*,[144] the English Court of Appeal recognised a scaling-down exercise as a proper response to clauses that overcompensate for loss.[145] Such scaling-down exercises would at least help the court to reach an approximation of the value of interest lost by the promisee, especially in cases where the loss is non-pecuniary nor challenging to estimate going by the compensatory principle in contract law. Thus, where remedial clauses are considered to overreach the proportionate quantum essential to protecting the promisee's legitimate interest in the light of facts available at the time of the breach, or the event it was to protect against, they should be read down to a reasonable quantum.

Conclusion

In the foregoing discussions, the thesis of this article was pursued with its pith being that courts should exercise more supervisory powers over remedial clauses. This should be done with a view to ensuring that compensation for losses resulting from contractual wrongs correspond to the loss suffered by a promisee, except in cases where the promisee can demonstrate legitimate interests deserving of super-compensatory protection. That argument was advanced based on three strands which may be recapitulated as follows:

- **a.** the courts should have wide powers to read down remedial clauses;
- **b.** the imperative to reconceptualise all remedial clauses, including termination clauses, under one single category; and
- **c.** the postulation and discussion of considerations that should guide courts in governing remedial clauses.

142 See *Lowe v Peers* (1768) 4 Burr 2225; *Winter v Trimmer* (1763) 1 W Bl 395; *Harrison v Wright* (1811) 13 East, 343.

143 See *Elsley v JG Collins Insurance Agencies* [1978] 2 SCR 916 (Canada).

144 [1989] 1 WLR 1026.

145 Ibid per Lord Nicholls, pages 1041–1042; see also the criticisms of the *Jobson v Johnson* decision in the UK Supreme Court in *Cavendish Square Holding BV v Talal El Makdessi* by Lords Neuberger and Sumption at paragraphs 84–87 and those of Lord Hodgson at paragraph 283.

The first strand of the recalibration agenda centres on the need to infuse the ideals of good faith and relational cooperation into contract law to accentuate private empowerment as the kernel of contract law. It is argued that the common law of contracts has always had ways of controlling contractual terms using default rules and standards, and expanding such control measures, particularly with regard to remedial clauses, is essential in our current economic milieu. Doing so will raise the bar of entitlement to agreed remedies clauses by ensuring that they do not overreach the protection of the legitimate interests of promisees. The imperative to factor society's interest in contract enforcement represents the nucleus of the first strand. The tenor of this strand is that contract law rules will serve and advance society's interest well where they do not allow remedial clauses to function as avenues for wealth extraction, upward redistribution of wealth that is undeserved, and the undue consolidation of economic power. If we adopt a libertarian attitude towards remedial clauses, even in business-to-business contexts, we will be creating room for social problems in the jejune belief of respecting freedom of contracts. That is particularly so in this age of yawning gaps in private law powers, facilitated, most notably, by the effects of commercial networks, market concentration and technology rights.

The possible counterarguments that a liberal attitude towards the revision of remedial clauses may heighten transaction costs and reduce economic certainty are weak and unsustainable. Equally weak are arguments that such a progressive approach galvanises paternalism in contract law. Adopting such an approach is not as intrusive as may be argued, even though it upsets the *status quo*. What it does is to re-allocate entitlements. Promisees may stipulate remedial clauses, but they should only be entitled to enforce such provisions if they can demonstrate a legitimate cause for doing so, particularly at the time of breach or the eventuation of an undesired event to which the provision caters. Such an approach would only reshape contract drafting, especially by making promisees identify with clarity what they consider to be legitimate interests. The legal adoption of such a standard does not amount to the removal of the promisee's right to contract for remedies; it only raises the bar of enforcement for such clauses. Also concerning paternalism, adopting the judicial attitude towards remedial provisions as canvassed for in this article is no more paternalistic than respecting the *status quo*.

The second and the third strands are complementary, and they build upon the first. The second strand drives home the argument that all remedial clauses are, in substance, same. That the categorisation of remedial clauses based on their ostensible differences is unnecessary and unhelpful and should be avoided. Instead, that all remedial clauses, including express termination clauses, should be treated as same and governed using uniform rules. The third strand postulates on considerations that should inform the legal attitude in the governance exercise. The postulated considerations are as follows:

1. an association with a breach should not be a precondition to the identification of a clause as remedial;
2. that a party seeking to enforce a remedial clause should demonstrate a legitimate commercial interest in enforcing the said clause;
3. that the demonstration of a legitimate interest in the enforcement of a remedial provision does not exempt judicial intervention where events *ex post* (i.e. at the time of enforcement) show that the commercial risks concerning which protection was sought did not turn out as grave as originally feared;
4. courts should take account of events happening at the time of breach or the contractually undesired event to adjust effects of the clause downwards or upwards.

5. recognition of a duty to mitigate against losses resulting or likely to result from a breach in cases of remedial provisions; and
6. the discrimination against purely contractual rights in the grant of relief against forfeiture is unsustainable.

NILQ 71(3): 457–479

Mediation in French administrative courts: what lessons for administrative justice?

SOPHIE BOYRON*

Birmingham Law School

Abstract

In 2016, the French Parliament introduced a new chapter on mediation in the Code of Administrative Justice. To succeed, this reform needs to reverse repeated failures in this field. In view of the significant challenge of embedding administrative mediation in the French administrative justice system, the reform and its implementation were informed by empirical findings arising from a mediation pilot set up by the administrative court of Grenoble in Spring 2013. An empirical study of the pilot and of the experience of rolling out administrative mediation in France forms the core of this article and the context in which to revisit foundational questions about mediation and administrative justice. I argue that mediation is not ill-suited to administrative law disputes, but that to be integrated in a system of administrative justice, mediation requires the negotiation of a dedicated environment triggering in turn the emergence of a pluralist administrative justice system.

Keywords: mediation; administrative justice; legal pluralism; French administrative law; legislative reform alternative dispute resolution.

Introduction

On 18 November 2016[1] the French Parliament adopted a wide-ranging reform, modernising the justice system for the twenty-first century. Among other initiatives, it introduced a new chapter on administrative mediation in the Code of Administrative Justice (CAJ). While French administrative judges have had 'a mission of conciliation' since 1986,[2] in reality, conciliation or mediation[3] has been a long time coming to the French administrative courts. The new legislation aimed to ensure that this method of dispute resolution finally took root. To succeed, the 2016 reform would have to reverse long-

* I am grateful to Fiona de Londras, Paul Craig, Lisa Webley, Richard Young and the two anonymous reviewers for their insightful comments on earlier drafts of this article. Any shortcomings remain my own.

1 See Law no 2016-1547 of 18 November 2016.

2 See former Article 211-4 of the CAJ: 'In the first instance administrative courts and in the administrative courts of appeal, the president of the court can, if the parties agree, set up a conciliation and appoint the person(s) charged with this mission.'

3 The terms *conciliation* and *mediation* are used interchangeably in this article. While there was a debate about their respective meanings in France, the 2016 reform uses the term 'mediation', effectively resolving the debate.

standing failures and dismantle the barriers that stand in the path of administrative mediation in France.

This is no modest task. Ten years ago, I studied the seemingly parallel rise of mediation in administrative law disputes in England, France and Germany[4] and concluded that mediation would never flourish in the French administrative justice system as then configured. This scepticism reflected known concerns of speed and costs, remaining obstacles to the development of administrative mediation, and a sense that the French administrative justice system was unsuited to its introduction. The question, therefore, was whether these concerns could be addressed and the mindset and systemic challenges that long stymied embrace of administrative mediation in France overcome.

Conscious of the significant challenge of embedding administrative mediation, the reform and its implementation reflected empirical findings arising from a mediation pilot set up by the administrative court of Grenoble. I was afforded a rare opportunity to study this ground-breaking pilot in the Grenoble court and to observe the process of implementing the 2016 reform. This enabled me to confront my earlier findings with recent empirical observations. Through the analysis of French administrative mediation reflecting on this pilot and presented in this article, I aim to increase understanding of administrative mediation more widely by drawing lessons where appropriate and to explore the conceptualisation of mediation for administrative justice.

Mediation is defined in the 2016 law as

> … a structured process, regardless of its name, by which two or more parties try to reach an agreement so as to solve their dispute amicably with the help of a third-party, the mediator, chosen by them or designated by the court with their consent.[5]

Administrative mediation refers to mediation in an administrative law setting in which at least one of the disputants is a public body.

While the literature on administrative mediation remains scant and often practice-oriented,[6] the notion of applying mediation to administrative law settings is challenging, with some suggestions that it is ill-suited to an administrative law context. There are a number of reasons for this. First, concerns regarding the speed and costs of administrative mediation and consistency of outcomes must be addressed. Second, the imbalance of power within an administrative dispute needs to be overcome. Third, the tensions between public law principles of accountability and transparency and the mediation principle of confidentiality need to be resolved. Fourth, the apparent incompatibility of adjudication and mediation as administrative justice mindsets and paradigms must be managed. However, until now there has been limited empirically grounded work that investigated these challenges and proposed remedies.

---

4 All three countries sought to embed court-led mediation in their administrative justice systems; as the organisation of administrative justice had otherwise little in common in all three systems, this case of 'spontaneous convergence' piqued my interest, see Sophie Boyron, 'The rise of mediation in administrative law disputes: experiences from England, France and Germany' [2006] Public Law 320; and Sophie Boyron, 'Mediation in administrative law: the identification of conflicting paradigms' [2007] European Public Law 263.

5 See Article L 213-1.

6 See, for instance, J-M Le Gars, 'La conciliation par le juge administratif' 2008 AJDA 1468; B Blohorn-Brenneur and J Biancarelli (eds), *Conciliation et mediation devant la juridiction administrative* (L'Harmattan 2015); J-M Le Gars, 'La juridiction administrative saisie par la médiation' 2016 AJDA 2272; A Minet-Leleu, 'La médiation administrative' 2017 RDP 1191; S Deygas, 'J 21: la médiation existe enfin en matière administrative' (2017) 2 (Février) Procédures Etude 8; in the UK, V Bondy and L Mulcahy (with M Doyle and V Reid), 'Mediation and judicial review: An empirical research study' (Public Law Project 2009).

This article and the empirical findings that underpin it challenge my earlier assessments of the likely success of administrative mediation in France. It offers an empirically informed perspective on how groundwork, organisation and leadership can create the conditions in which administrative mediation can successfully be embraced. More broadly, it takes the experience of rolling out administrative mediation in France as the context in which to revisit foundational questions about mediation and administrative justice. I argue first that mediation is not ill-suited to administrative law disputes, second that mediation and adjudication can coexist despite their significant differences, and third that to be integrated in a system of administrative justice, mediation requires the negotiation of a dedicated environment triggering in turn the emergence of a pluralist administrative justice system. If that is the case, orthodox accounts of administrative justice are inadequate in light of this emerging reality, and the need to recognise and theorise the pluralism of administrative justice is pressing.

The article is organised in five parts. First, I provide a concise explanation of the challenging context and history of attempting to introduce administrative mediation in France (part 1), after which I briefly outline the methodological approach adopted in the study (part 2), and then, the Grenoble pilot and what it told us about the possibility of administrative mediation (part 3). Part 4 outlines the 2016 law reform, showing how the lessons from Grenoble reflected its formation and informed its implementation in order to overcome the difficulties identified in part 1. Finally, building on the French experience, part 5 will expand the existing theoretical understanding of administrative mediation and the challenges it brings to a system of administrative justice.

1 The introduction of administrative mediation in France: a long history of failure

France has repeatedly tried (and failed) to introduce administrative mediation. This long history of failure starts with Article 22 of the law of 6 January 1986, which recognised a 'mission of conciliation' in first instance administrative courts.[7] While this law refers to conciliation, efforts moved to mediation – a more widely known dispute resolution mechanism. Over the next 30 years, various initiatives were adopted to promote the use of administrative mediation/conciliation, however, they were sporadic, irregular and lacked follow-up.[8] In 1993, the Conseil d'Etat produced a report entitled 'Solving disputes differently'[9] that promoted the use of alternative dispute resolution (ADR) in the administrative courts and the administration, leading to the adoption in 1995 of a circular 'informing' the civil service and all public bodies of the availability of settlements when solving disputes.[10]

In 1999, due to a lack of progress, a working group chaired by President Labetoulle, a senior member of the Conseil d'Etat, aimed to encourage the use of conciliation in the administrative courts. The committee drafted a set of informal guidelines but judged it unnecessary to adopt any formal implementation. Consequently, obstacles such as the

7 See Article 22 of law no 14-86 of 6 January 1986: 'The administrative courts exercise also a mission of conciliation.'

8 For instance, see J-M Le Gars, ' La conciliation par le juge administratif' AJDA 2008, 1468; and J-M Le Gars and S Wegner ,'Evolution de la médiation et de la conciliation devant la juridiction administrative française' in B Blohorn-Brenneur (ed.), *Les nouveaux enjeux de la médiation* (L'Harmattan 2016).

9 *Résoudre autrement les conflits: conciliation, transaction, arbitrage en matière administrative* – Rapport du Conseil d'Etat (1993).

10 Circular of 6 February 1995 concerning the development of settlements to solve disputes amicably, JO 15 February 1995, 2518.

two-month time-limit for judicial review, the financial costs of mediation, and the bindingness of mediation agreements were never addressed, contributing to a lack of concrete progress. Finally, in 2002, at the request of the administrative court of Melun, the Conseil d'Etat recognised the practice of 'certification' of conciliation agreements, thereby granting them the same legal effect as court judgments.[11] This meant that one possible hurdle – enforceability and bindingness – was overcome, but there was nevertheless a distinct lack of progress. This only intensified in the period between 2002 and 2008, when momentum seemed to be lost and there was a dearth of initiatives oriented towards the adoption of administrative mediation.

However, EU law reignited matters in 2008, with the adoption of the Directive on Certain Aspects of Mediation in Civil and Commercial Matters.[12] While the Directive required that mediation be offered in the large majority of cross-border disputes, it aimed to promote access to mediation more generally. In response, the Conseil d'Etat published a report on 'developing mediation in the context of the European Union',[13] and the Directive was transposed into domestic law by the Ordnance of 16 November 2011. This Ordnance introduced mediation for trans-border disputes in the CAJ, finally to at least some degree lodging administrative mediation in the relevant domestic law. In December 2011, the mission of conciliation was extended to administrative courts of appeal,[14] and a further circular (replacing that of 1995) 'reminded' the administration that settlements should be used to prevent and solve disputes whenever possible.[15]

The 2008 Directive, thus, did lead to a limited doctrinal embrace of administrative mediation,[16] but systemic challenges to its full embrace remained. Even after 30 years of commitment to conciliation there was no formal implementation, judges had never received any mediation or conciliation training, and the various 'initiatives' failed to embed a practice of conciliation/mediation in the administrative courts. This was the context in which commentators, including me, had developed a scepticism about the likely success of administrative mediation in France.

That scepticism was only exacerbated by an awareness of the unsuitability of the French administrative justice system to the introduction of administrative mediation (and ADR more generally). Although, by 1986, administrative courts had developed an effective system of review of administrative decisions, administrative justice was remote from the citizens it aimed to serve. While administrative courts applied rigorous scrutiny to the question of the legality of administrative acts, remedies for illegality were limited to quashing decisions,[17] delivering a rather abstract and often delayed kind of administrative justice.[18] Also, administrative courts were victims of their own success: they were struggling with a spiralling caseload and long delays in hearing cases. Furthermore, more claimants were complaining to the courts that judgments in their

11 CE Ass 6 December 2002 Advice no 249153 Haÿ-les-Roses.

12 2008/52/EC of 21 May 2008.

13 *Développer la médiation dans le cadre de l'Union européenne* – Rapport du Conseil d'Etat (2010).

14 Law no 2011-1862 of 13 December 2011.

15 Circular of 6 April 2011 concerning the development of settlements to solve disputes amicably, JO 8 April 2011.

16 Mediation has never been used to solve a cross-border dispute as yet.

17 This is called *Recours pour Excès de Pouvoir* (or review for excess of power) and it is the functional equivalent of judicial review.

18 In the iconic *Benjamin* case regarding the ban of a literary conference (CE 19 May 1933), the Conseil d'Etat proclaimed the fundamental character of freedom of assembly, applied an 'anxious scrutiny' and vindicated Benjamin but handed down the decision three years after the ban.

favour remained unexecuted. Finally, no thought had been given to how mediation might be pursued in a system of inquisitorial and written procedures. In this context it seemed optimistic, at best, to think that administrative courts could and would adopt a conciliation mission and have the resources and capacity to implement it within their already strained circumstances.

However, the situation has changed. First, while the caseload has continued to grow,[19] administrative courts are now more efficient and better resourced, and the average time for a judgment has been noticeably shortened.[20] Reforms have given administrative courts tools to intervene in urgent circumstances,[21] diversified the remedies available to judges,[22] and tried to ensure a more timely execution of judgments.[23] Administrative justice has shifted its focus to embrace better the reality of claimants' experiences. By the time of the 2016 reform, the administrative justice system was arguably in a better position than it had ever been to fulfil the conciliation mission that had so long remained somewhat notional.

2 Methodology

This study of administrative mediation in France is based on my observation of the mediation pilot, set up in spring 2013 by Vice-President Wegner of the *tribunal administratif* of Grenoble,[24] and on my participation in the implementation of the 2016 reform.

2.1 The Grenoble mediation pilot

Having established the pilot in Grenoble in spring 2013, Vice-President Wegner engaged with me in Summer 2016 to co-design an observation study, which involved elements both inside and outside of the court. The study was designed to help me develop an opinion, but not a formal evaluation, of the pilot from which guidance might be developed on the nature of disputes suited to mediation. For my part, I was keen to investigate this rare pilot, particularly given my earlier work. In designing the study, we were influenced by methodological and fieldwork accounts taken from anthropology and organisational research with the work of Bruno Latour,[25] a notable inspiration.[26]

Within the court, the study combined two months of participant observation (December 2016–January 2017) with complete access to all 45 mediation court files and

19 In 2016, 234,460 new cases were introduced in the administrative courts. This needs to be contrasted with the figures for 2006 (170,000), 1991(70,000) and 1976 (20,000).

20 In the last decade, the average time for a judgment decreased from two years to 16 months in the administrative courts, from three years to 14 months in the administrative courts of appeal and from 14 months to less than 10 months in the Conseil d'Etat.

21 Since 2001, administrative courts can issue *référés* (interlocutory injunctions): the *référé-liberté* orders the administration to act/cease to act when a decision violates a claimant's fundamental right in a serious and manifestly illegal manner (Article L521-2 CAJ) and the *référé-suspension* stays an administrative decision for reasons of urgency when there are serious grounds for believing the decision to be illegal (Article L521-1 CAJ).

22 This evolution results from a combination of legislation and case law, see F Blanco, *Pouvoirs du juge et contentieux administratif de la légalité: contribution à l'étude de l'évolution et du renouveau des techniques juridictionnelles dans le contentieux de l'excès de pouvoir* (PUAM 2010).

23 A law of 8 February 1995 gave administrative courts the power to issue a 'preventive' injunction detailing the steps for enforcement or to issue an injunction once a judgment's execution is contested or delayed.

24 Wegner was one of eight Vice-Presidents in the Grenoble court. Each Vice-President heads a chamber of the court.

25 Bruno Latour, *La fabrique du droit* (Poche 2009).

26 See also B Morean, *Ethnography at Work* (Berg 2006) and T Eriksen, *Small Places, Large Issues: An Introduction to Social and Cultural Anthropology* (Pluto Press 2015).

full access to all activities relevant to the mediation pilot. Following the observation period, I held a feedback session at which my preliminary findings were presented to and discussed with judges and clerks of the Grenoble court.

Outside the court, the study comprised three elements: participation in a one-day mediation training workshop attended by all of the mediators acting for the administrative court; observation of the meetings of all live mediations over the period of observation (three in total); and semi-structured interviews with key actors in the mediation pilot – namely, mediators regularly appointed by the court (6), public law barristers who had either proposed a mediation to the court or represented a party in mediation (5), and civil servants in local authorities' legal or human resources departments who had been involved in at least one mediation (5) and, finally, the coordinator of the Centre for Mediation, the court's main mediation provider.

During this period, I was careful to limit any direct participation in the pilot's activities. However, towards the end, I organised a survey of all new cases of one chamber with a view to trialling my list of indicators for the identification of mediation disputes. I recommended mediation in those cases I judged suitable. Upon my departure, the few that met with the approval of the chamber's president were forwarded to the newly appointed mediation champion for his decision.

2.2 The implementation of the 2016 reform

In January 2017, I was invited to sit on the National Steering Committee on Administrative Justice and Mediation, which was the body charged with securing the implementation of the 2016 reform. The committee's president was interested in the Grenoble pilot and the identification of mediation cases, both of which I had insight on following my time observing the pilot. For myself, this was an opportunity to observe (and possibly inform) the implementation of a key reform. All but one of the committee meetings took place after the observation period in Grenoble. From February 2017, my participation in the meetings was informed by the practice of observant-participation.[27] Importantly, I only ever joined in the discussions and argued points on the basis of my findings, understanding and experience of the Grenoble pilot.

3 The Grenoble pilot

To understand the 2016 reform, it is important to begin with the mediation pilot. This pilot played a key role in the reform process: not only are its choices reflected in the law of 18 November 2016, but it trialled many solutions reproduced in the implementing legislation. In this section, I present an account of the Grenoble pilot, reflecting my experience of observing the pilot and, subsequently, participating in the implementation of the 2016 reform.

3.1 The structure of the pilot

To an outsider, the choice of Grenoble as a place to trial administrative mediation may seem random. However, mediation has been relied upon in Grenoble's private law courts since the mid-1990s. This meant that Grenoble has a local culture of mediation and a large pool of professional mediators available for the pilot. Vice-President Wegner was

27 This concept was coined by B Morean, 'From participant-observation to observant-participation: anthropology, fieldwork and organizational ethnography' Creative Encounters Working Paper No 2 (Copenhagen Business School, July 2007). If the opportunity arises, he advocates increasing participation in the activities of the community or organisation under observation, to move from observant-participation to participant-observation for better access to and truer interactions with people.

personally responsible for introducing and running the court-led pilot. Before the pilot's launch, the ground was carefully prepared outside the court by negotiating a formal agreement with key stakeholders: the court, the town of Grenoble, the *département* of Isère, Isère's centre for local government's civil service, the Bar of Grenoble, and the Centre for Mediation of the Chamber of Commerce and Industry. This agreement stated that it aimed to implement the mission of conciliation of the CAJ[28] and listed the substantive areas likely to attract a resolution via mediation (e.g. public employment, planning, public procurement, social welfare). In addition, it removed many of the practical obstacles to the development of administrative mediation by:

- indicating that mediation would be triggered by agreement from both parties after proposal by one party or a judge;
- allowing the suspension of the two-month time-limit;
- regulating the duration – three months renewable once;
- granting the option of 'certification' by the court of a mediation agreement;
- regulating mediation costs – equal share of mediation costs between all disputants;
- safeguarding both impartiality and confidentiality; and
- adopting an ethics charter.

The conclusion of the agreement gave the opportunity to begin acculturating key actors – including from the public sector – as well as to ensure buy-in from all stakeholders.

In France, introducing an action in the first instance administrative courts is free, does not require a barrister (although claimants usually instruct one) and legal aid is available, thereby creating financial incentive to go to court rather than mediation. Consequently, Vice-President Wegner tried to address the question of mediation costs: a €500 ceiling price per mediation was negotiated with the Centre for Mediation and free access to meeting rooms. Furthermore, the Grenoble court has two judges, both trained mediators, who acted for free when a disputant could not afford the mediation costs.[29] Finally, in employment disputes, public authorities were 'encouraged' to cover the mediator's fee in full.[30] Even so, mediation was not 'the cheapest option' for some disputants (especially if they had not planned to instruct a barrister for their court action).

This careful set-up contrasted markedly with the lack of organisation inside the court. There, responsibility and day-to-day running of the pilot rested entirely with Vice-President Wegner. No clerk was allocated to this new function and no process was put in place for the identification of mediation cases. While the pilot had the support of the court's president, the extra work was simply shouldered by Vice-President Wegner without workload adjustment. This had consequences for the development of the pilot.

3.2 Identifying disputes for mediation

Vice-President Wegner played a pivotal role at the start of most mediations during the pilot by identifying suitable disputes and choosing the mediator.

With no internal process, mediation disputes were identified in an artisanal fashion: Vice-President Wegner reviewed at irregular intervals the case lists of chambers that dealt

28 See former L 211-4 CAJ.

29 If the mediation were to fail, the judge-mediator would not be allowed to work on the case after it returned to court.

30 Employment disputes are covered by the experiment in compulsory mediation, and consequently mediation is now free for those disputes.

with public contracts, public employment disputes, public works etc. In addition, colleagues would signal to him cases they thought might be suitable early in the court's process. After perusing the file and discussing it with the judge responsible for the case, Vice-President Wegner would decide whether to propose mediation; if so, approval of the chamber's president was sought. For this to work, all interlocutors needed to be convinced of the benefit of mediation. As support for the pilot varied with individual judges, this identification was neither uniform nor systematic. Vice-President Wegner was not in a position 'to catch' all potential mediation disputes and would never have had the capacity to handle them. While the pilot was well-known in the court due to the seniority and standing of Vice-President Wegner, it was generally seen as sitting at the margins of the court's work due to the small numbers of cases involved.

Vice-President Wegner (and his colleagues) were driven by the belief (which itself motivated the pilot) that some disputes are better resolved through mediation. This mediation mantra drove the pilot and guided the identification of mediation disputes. In addition, and recognising mediation as a voluntary process, the court tended to propose mediation (if appropriate) when one or both parties requested it.

Analysing the features of Grenoble's past mediation cases, at the end of the observation period I developed a list of indicators that systematised Vice-President Wegner's practice in identifying cases for mediation. These indicators were not applied during the observation period or to the pilot as a whole, and thus did not distort the existing practice of the court. These were that the dispute:

1. has a degree of urgency;
2. requires a bespoke solution (beyond the range of judicial remedies);
3. is wider or the interests more numerous that the one(s) in litigation, so that the judicial process would not resolve the real dispute or reach all disputants;
4. has complex facts with mediation being a better forum for their exploration than the court's written procedure;
5. is emotionally charged (with the judicial process unable to pacify them);
6. concerns disputants with an ongoing relationship that may be compromised by the judicial process; and
7. concerns disputants that are both public bodies (often a symptom of administrative dysfunction beyond the reach of the judicial process).

A review of the practice also suggested that some features tended against mediation. I identified these as follows:

1. (psychological or mental) vulnerability of a disputant;
2. the need to decide a point of law;
3. public order considerations; and
4. the possible manipulation of the mediation by one or both disputants for harmful aims

3.3 Choosing the mediator

Appointing the mediators was the responsibility of Vice-President Wegner. He knew (and had sometimes trained with) the mediators and took care to appoint whom he considered to be the 'right' mediator for a given dispute, often after discussion with the coordinator of the Centre for Mediation. All mediators were fully qualified with years of mediation experience, but Vice-President Wegner chose individual mediators for their legal expertise

and experience in the substantive area of the dispute (public procurement, public works, public employment etc.). This enabled mediators to rely on their knowledge of the legal context, policy background and administrative landscape to strengthen their facilitative capabilities.

The court chose to appoint external mediators in 80 per cent of mediation cases. For the remaining 20 per cent, one of two in-house mediators was appointed to reduce costs for the parties or to reassure the parties in sensitive disputes.

3.4 The practice of mediation in the Grenoble pilot

From my observation and interviews it became clear that through the Grenoble pilot a distinct practice of mediation had developed in respect of type process, and duration of mediation. I consider all three of these in turn.

3.4.1 Type of mediation

Although evaluative mediation tends to be favoured[31] in public law settings, mediators practised facilitative mediation with quasi-systematic reliance on joint meetings. Sometimes called 'pure mediation', this concentrates on the needs and interests of the parties, favours problem-solving and limits the mediator's role by making the parties (with their legal representatives) responsible for finding a solution.[32] In Grenoble, the use of facilitative mediation reflected a genuine attempt to move away from the limitations of legal discourse and strictures of inquisitorial judicial procedure. From my observations and interviews, it became clear that, while mediators were careful to avoid suggesting solutions, they took an active role in shaping the resolution process, through restatements, judicious questioning and making demands.

The use of facilitative mediation arguably increases the risk of power imbalance in administrative mediation. However, it was clear from my observation and interviews that mediators were alive to this issue and sought to address it in several ways. First, they ensured that individual disputants had the time to explain the problem(s), present their demand(s) and formulate solutions. Second, mediators insisted on constructive participation of both disputants. For example, I witnessed mediators challenging the choice of the authority's representative when this threatened to impede progress (e.g. because of the representative's lack of authority); in the pilot, it was standard practice for a line manager or elected representative (e.g. mayor) to represent the public authority. I have also observed mediators (successfully) demand that public authorities brought viable solutions to the table. Third, mediators relied on legal representatives to generate a more level playing-field between the disputants. For the great majority of mediations I surveyed or observed in Grenoble, disputants were legally represented at the mediation meetings. By supporting their clients throughout the process, particularly when formulating solutions or advising on proposals, barristers helped mediators to create a space for *bona fide* problem-solving. Fourth, mediators' practice of supporting the consultation of other public bodies or experts when necessary was often instrumental in bringing knowledge of possible solutions to the disputants, which tended to empower private disputants and further level the playing field.

---

31 Evaluative mediation tends to focus on the legal entitlements and rights of parties, the mediator playing an active role and bringing to bear her substantive expertise to provide advice, evaluate the claims and propose a solution.

32 See L Boulle and M Nesic, *Mediation: Principles, Process and Practices* (Butterworths 2001) 27.

3.4.2 Process of mediation

In the Grenoble pilot, the search for a solution shaped the organisation of mediation meetings, the choice of participants and the consultation of other public bodies or experts.[33] This freedom in designing the mediation and its process is one main advantage. All 'interested parties' can be invited to participate, not just the parties in the case. For example, one case involved a local authority that had contracted a public works corporation to restore the confluence of two rivers. The work caused the adjoining road to subside and was stopped while repairs were undertaken. Then soon after the contracted work had been completed an exceptional flood mostly destroyed it. In this case, the contractor, sub-contractors, project manager and local authority all participated in the mediation, even though sub-contractors could never have been a party in the administrative court case.[34] Flexibility in process thus allowed a fuller dispute resolution to be pursued here.

Flexibility also allowed for public participation to be fostered. This is well illustrated by another case, which involved a sale of land from a local government. This was blocked because of strong opposition in the village, despite the buyer having fulfilled all the pre-conditions and acquired the right to buy. The case involved strong emotions among inhabitants and elected representatives. When the local council rejected the first mediation agreement, the mediation was opened to those who opposed the sale, even though they were not disputants in the case. This resulted in frank discussions between the opposition and the prospective buyer, the redrafting of land-use restrictions in the mediation agreement and, ultimately, the sale being completed. Here, mediation complemented the local government's decision-making process and gave a voice to the local community.

The Grenoble pilot also revealed the role that can be played by other public bodies when canvassing the viability of a solution in mediation. In one mediation concerned with damage caused in part by public works to a private but listed property, the *Bâtiments de France*'s[35] advice that the repairs need not use the same (expensive) traditional method as had been used prior to the damage helped to lower the cost of repairs and for the parties to reach a financially viable agreement through mediation.

3.4.3 Duration of mediation

Mediation is commonly presented as a time-saving device. However, the pilot demonstrated that administrative mediation is not particularly speedy. On average, it took nine months for parties in a successful mediation to reach an agreement and withdraw their action. While the 2013 agreement imposed a limit of six months, in reality, Vice-President Wegner renewed on request while progress was being made. This flexibility in terms of duration proved to be important. It became very clear that bespoke solutions (discussed below) take time to be formulated and for their acceptability to be explored. In addition, there was often a significant period between signing the agreement and withdrawal of the action, explained mostly by the need for the agreement to be endorsed formally by the public authority. While the latter mode of endorsement is French-specific and unlikely to happen elsewhere, with an average of 18 months for a judgment in Grenoble, mediation is still quicker than the judicial process and carries no risk of appeal.

33 The choice of participants is often key to success, see for instance, ibid 120.

34 Because of the French public–private divide, all disputes involving sub-contractors come under the jurisdiction of the private law courts.

35 The *Bâtiments de France* is the public authority of state architects responsible for urban conservation and heritage.

3.5 The mediation agreements

From my observations, interviews and survey of the pilot, it became clear that the mediation agreements arrived at in the Grenoble pilot aimed to frame bespoke but legal solutions.

3.5.1 The outcomes

In the Grenoble pilot, many parties chose mediation to design bespoke solutions. In 60 per cent of successful mediations, the agreement contained a solution that would not have been available (easily or at all) through the administrative court process. To give a flavour of these, four disputes are reproduced below with their outcomes:

Dispute 1: a dispute about the fee structure of a public domain concession contract for the management of a major regional transport infrastructure triggers the introduction in rapid succession of six court actions while the contract has still 14 years to run. The managing company argued that the fee calculations rested on inaccurate forecasting of costs and profits. The dispute threatened the continuity of the contract, despite this transport infrastructure being central to the regional economy. The mediation resulted in the following: cancellation of the 2010 payment for the public service contribution fees and cancellation of the public domain occupancy fees for the first two years of the contract. In addition, the contractual provision containing the fee calculation was re-drafted and applied to year three and for the future.

Dispute 2: the sale of land from a local government was blocked by a local authority because of strong opposition in the village, despite the buyer (who was originally from the village) having fulfilled all the pre-conditions and acquired the right to buy. A first mediation between the buyer and the representative of the local authority led to an agreement. However, the council of the local authority, when asked to endorse it, altered the drafting of the strict easements and use of land restrictions. An additional mediation meeting was organised with the buyer, the representative of the local government and those opposing the sale. After a frank discussion that helped clear several misunderstandings, a new agreement was arrived at with easements and use of land restrictions redrafted to address local concerns. The sale was then approved by the local authority and completed.

Dispute 3: the Rectorate,[36] having committed itself to 'lending' a music teacher to a new local public authority for people living with disability, later 'forgot' to do so through loss of institutional memory, forcing the authority to employ a costly replacement. In addition to solving the immediate dispute – compensation for employing a replacement – the mediation allowed the creation of processes and the identification of contacts in both authorities to avoid recurrence.

Dispute 4: A local authority contracted a public works corporation to restore the confluence of two rivers. The work caused the adjoining road to subside and was stopped while repairs were undertaken. Then soon after the contracted work had been completed an exceptional flood mostly destroyed it. The contractor, sub-contractors, project manager and local authority were all in dispute over the repairs, the various liabilities and the share of the costs. However, according to French administrative law sub-contractors cannot be a party in an action before the administrative courts. Consequently, there was a danger that the court would fail to solve the dispute for all. The mediation allowed all disputants to agree on their individual share of liability, the repairs needed and the apportionment of costs.

36 The Rectorate is responsible of all public education (from primary school to university) in the academic region and ensures the implementation of the Ministry of Education's policy in the academic region.

The presence of legal professionals in the mediations I observed did not result in the 'law-ification' of the debates: legal claims were not made to justify demands (or their rejection). Instead, all present concentrated on solving the root cause of the dispute. Notwithstanding this, however bespoke the solution, my observations and interviews reveal a clear understanding among mediators, barristers and public authority representatives that the mediation agreement needed to be legal. Importantly, the ethics charter attached to the 2013 Agreement describes the role and obligations of mediators and specifies in article 11 that:

> [the mediator] acts in compliance with the law: he/she immediately reminds the parties that any proposal that does not respect public order or the interests of concerned third parties causes the mediation to be stopped immediately.

From evidence on file, Vice-President Wegner was sometimes consulted by mediators on behalf of both disputants as to the legality of specific solutions. While Vice-President Wegner worried that his advice might compromise the mediation process, this practice speaks to the concern about legality. Consultation of other public bodies or external experts was often aimed at eliciting advice on the acceptability (including legality) of solutions or about 'legal' alternatives. The fact that the pilot was court-led and that key actors in mediation are legal professionals or legally trained (e.g. civil servants) goes some way to explaining this common commitment to legality and the rule of law. Furthermore, the legality of mediation agreements can be challenged in the administrative courts. Often the process of certification is used to pre-empt a legal challenge.

3.5.2 The certification of mediation agreements

The administrative court of Grenoble only required to be informed that an agreement had been reached, not of its content. Vice-President Wegner wanted to protect confidentiality and highlight the independence of the mediation process from the court. An agreement's content was only divulged to the court when disputants asked for certification. The court tended to discourage this: Vice-President Wegner worried that focusing on certification would compromise the search for a creative solution. He also wanted to avoid legitimising mediation through the judicial process and increasing the court's workload. Five certifications (11%) were sought during the pilot and all were granted (e.g. the mediation agreement of dispute 1).

3.5.3 Compliance

No problem of compliance with a mediation agreement arose during the pilot, a notable contrast from the problems over execution of court judgments mentioned previously.

3.6 A RELATIVE SUCCESS

Set up in spring 2013, the pilot operated for four years until it was overtaken by the new legislation. While defining 'success' in mediation is a minefield, the court settled on the withdrawal of the action. By that definition the pilot might be described as (very) small but successful. Over the four-year period of the pilot, there were 45 mediations (covering 58 disputes),[37] with 75 per cent of the associated actions being withdrawn. In all but one,[38] a mediation agreement was signed between the disputants. For the remaining 25 per cent, the case simply continued in the court. This success rate may seem an achievement, but an average of 11 mediations per year is (very) little in view of the 8000

37 Some mediations brought together multiple court actions or disputes.

38 In this case, the claimant just withdrew the court case.

new cases introduced in the Grenoble court every year. This very low figure is largely the product of the pilot's set-up: Vice-President Wegner had little spare capacity.

Furthermore, over the same period, public bodies rejected outright approximately 50 per cent of all mediation proposals. This rejection rate is troublesome, particularly since Vice-President Wegner sometimes attempted to talk the public authority into accepting mediation when a private party had already done so. This suggests that further work is needed to shift the mindset of public authorities to accept mediation when proposed rather than pushing for a court judgment.

However, overall the pilot suggested that administrative mediation *can* work and, when an agreement between the disputants is reached, that it is an effective method of dispute resolution including in French administrative context. Importantly, neither the pilot nor the 2016 reform made claims of time or cost savings. In fact, administrative mediation may be quicker than a judgment, but it takes longer than anticipated. Even though steps were taken to minimise financial costs for the disputants, mediation is more costly than a court action for some disputants. Furthermore, administrative mediation can give rise to worries of inconsistent treatment between those individuals benefiting from a mediation agreement and those who do not. However, the particularity of disputes and of the solutions agreed upon mean that (in)consistency may not be a matter for concern. In reality, these bespoke outcomes were not transferable to other disputes. Inconsistency could be become a concern if the basis for identification of mediation disputes changed. Furthermore, granting an advantage to a disputant to the exclusion of others in the same situation would constitute a clear breach of the French constitutional principle of equality, a principle firmly protected by all French courts.

Still, the full compliance from public authorities where they participate in the mediation is significant.

Critically, the pilot confirmed that administrative mediation needs formal implementation to remove the obstacles to its development and will be aided by the commitment of a 'mediation champion' within the court. This suggests that successful implementation of administrative mediation requires not only legislation but also committed leadership so that mediation grows. Overall, the Grenoble pilot, then, pushed back against assumptions from previous research as well as the history of failure outlined in part 1 to underpin the claim that it was realistic to move towards implementing mediation on a national basis. In this respect, the 2016 law reform was a momentous development.

4 2016: A ground-breaking reform

The 2016 law added a chapter to the CAJ with the aim of regulating administrative mediation; it also established an experiment for compulsory mediation. The new chapter is divided into three sections: the first and general section contains a standard definition of mediation,[39] lists the qualities that a mediator must possess (impartiality, competence and diligence), regulates confidentiality and recognises the possibility for the certification of a mediation agreement. The other two sections regulate separately mediation triggered by the parties outside of any legal action and mediation triggered by a judge (on request or approval of the parties) once an action is introduced in the administrative court. Both sections set down detailed rules that aim to remove previous obstacles to mediation.

39 See Article 213-1 CAJ.

4.1 The new law

The new legislation mirrors many choices made by the Grenoble pilot and, consequently, removes the previous obstacles to administrative mediation in France. In particular, it suspends the two-month time limit for triggering judicial review, recognised as a major obstacle to mediation. Second, it allows for legal aid to cover mediation costs where mediation is triggered by a judge (compulsory mediation is free). It also allows for *pro bono* work by mediators. Removing the time and cost barriers, then, enabled parties to opt for mediation in a reasonably low-risk way. Third, the law allows for the certification of mediation agreements, thus strengthening their legal effect where desired by parties. All of these innovations, then, reflect the experience in Grenoble to address systemic barriers to administrative mediation.

In addition to this, Parliament introduced an experiment for compulsory mediation for social welfare and public employment disputes. In essence, these are the most disputatious areas of French public administration, with the possibility that these types of cases will create a rich dataset for further study and provide an impetus for the deployment of mediation across the public sector.

4.2 Implementing the new law: Grenoble's influence

The secondary legislation required to give effect to the new law was adopted on 18 April 2017[40] and the reform came into effect on 1 January 2018. In the period between the passage of the new law and its coming into effect, close attention was paid to the results of the Grenoble pilot. The Conseil d'Etat established a National Steering Committee on Administrative Justice and Mediation (NSCAJM), chaired by President Libert, honorary Administrative Court President. The administrative court of Grenoble was well represented on the NSCAJM with its two judge–mediators and later myself being members. This presence allowed for the Grenoble experience to be brought to the Committee and proved decisive in respect of four issues: choice of mediators, duration of mediation, role of the court and identification of the mediation disputes.

As outlined in part 3, the practice in Grenoble was to appoint mediators who are both experienced in the role and experts in the subject-matter of the dispute. This 'double qualification' is not standard mediation practice, particularly when practising facilitative mediation. However, new Article R 213-3 CAJ reproduces this double requirement, clearly influenced by its successful deployment in Grenoble. More controversial was the decision to appoint administrative judges as in-house mediators, as was the practice in Grenoble. Concerns were expressed in the NSCAJM about administrative judges' lack of formal mediation training: some members felt that mediation should be entirely 'externalised', while others argued that it would be counter-productive to side-line those judges with knowledge of mediation, as they could play a key role in implementing the reform in their court. In the end, new Article R 213-3 CAJ requires all mediators to have mediation training or experience, thus enabling judges to act as in-house mediators where they have the required qualification.

With regard to the issue of duration, the original draft reproduced the provision contained in Grenoble's 2013 Agreement and imposed a three-month time limit renewable once. However, in practice the time limit was frequently extended beyond six

40 Decree no 2017-566 of 18 April 2017 concerning mediation in the disputes coming under the jurisdiction of the administrative judge.

months. So, learning from the pilot, the final text imposed no time limit but left the duration of mediation to the discretion of the court.[41]

The regulation makes it clear that, throughout the mediation process, the overall responsibility for the dispute remains with the court.[42] For instance, according to Article R 213-8 CAJ, the court can order interlocutory measures during this period. In addition, the regulation stipulates that the mediator can inform the court of any difficulties encountered during the mediation, allowing the mediator to seek advice or simply to discuss a problem. Again, this matches the practice in Grenoble where Vice-President Wegner was always available to discuss difficulties,[43] but mediators acted with complete freedom and independence.

The identification of mediation disputes is critical to the successful deployment of administrative mediation. Still, developing a means of identifying such disputes is anything but straightforward. For this reason, the NSCAJM adopted the list of indicators trialled in Grenoble, as discussed in part 2 above. The list was included in the national judicial guideline and now guides identification of cases for mediation throughout the country.[44]

Building on the legislative measures and judicial guidance and drawing upon the experience of the Agreement and groundwork in Grenoble, national and local implementation strategies were developed to maximise the adoption and practice of administrative mediation across the country. The NSCAJM instigated the compilation of a national list of mediators for use in the administrative courts and organised a training programme of sensitisation to mediation aimed largely at administrative judges. In between committee meetings, President Libert toured the country and the administrative courts to present the new legislation, smoothing the path of the reform with the relevant government departments and administrative authorities, setting up a network of partners for the compulsory mediation experiment, and meeting a wide range of stakeholders to ensure their support. Thus, President Libert adopted a systematic strategy of information and consultation with national stakeholders to ensure the successful implementation and acclimatisation of administrative mediation. In addition, the Conseil d'Etat signed, on 13 December 2017, a national framework agreement with the National Bar Council that promoted the use of administrative mediation among legal professionals. Finally, to ensure that mediation became a reality in all of the country's 42 first instance administrative courts, the Conseil d'Etat encouraged the appointment in each court of mediation champions, mirroring closely the approach of Vice-President Wegner in Grenoble, and created an achievable target for mediation cases (1% of the annual caseload of each court for 2019, i.e. a total of 80 for the Grenoble court).

Following on from its pilot, Grenoble revisited its pre-pilot strategy and expanded it, signing Agreements with the Bars of the three remaining *départements* under its jurisdiction in order to be able to recommend mediation whatever the claimant's place of residence. Furthermore, the President of the court and the mediation champion organised a series of meetings with the legal departments of several public bodies within the court's jurisdiction to present the reform, signal the court's support and discuss ways of growing mediation together. Consequently, new relationships, working practices and pathways were being forged to strengthen the local readiness for administrative mediation. Inside

41 See Article R 213-8 CAJ.

42 See Article R 213-5 CAJ.

43 From evidence on file, this happened albeit exceptionally.

44 See *Vademecum de la Médiation* – Conseil d'Etat (June 2017).

the court, one of the two judge–mediators was appointed mediation champion following the departure of Vice-President Wegner. New pathways have been set-up to ensure a better identification of mediation disputes and a clerk has been appointed to support the work of the mediation champion. Finally, to ensure that mediation takes its rightful place in the practices of the court, the President is contemplating a new staff performance target focused on mediation. Thus, Grenoble continues to develop its local strategy to maximise the implementation of the new law, building on its local knowledge, relationships and context and ensuring receptiveness inside and outside of the court. Early indications show that the large majority of first instance administrative courts are deploying a similar local strategy.[45]

5 Revisiting mediation in a system of administrative justice: early reflections from the French experience

With the formal implementation of the new legislation now completed, it is time to draw some early conclusions arising from the French experience. Albeit limited in scope, this study of administrative mediation has four main implications for the wider system of administrative justice: first, administrative mediation reflects and is shaped by the specificities of the administrative law setting; second, in practice administrative mediation finds ways of accommodating the conflicting paradigms to which mediation and adjudication belong; third, administrative mediation seems to require the creation of a dedicated (sub)-system of informal administrative justice; and, fourth, this facilitates the emergence of a plural system of administrative justice.

5.1 The specificity of administrative mediation

The French experience helps to understand how mediation is transformed by its contact with administrative law. Indeed, for administrative mediation to be a 'viable option' in any system,[46] it is important to understand that the specificity of administrative justice shapes this species of mediation. In other words, administrative mediation needs appropriate adjustments to account for the nature of administrative law and public administration.

5.1.1 Flexible decision-making process

From the survey and observation of the pilot's mediation disputes, it seems to me that administrative mediation blurs the distinction between remedy and decision-making. While the overall structure (two disputants or more resolving their dispute with the help of an impartial third-party) is definitely 'remedial' in nature, administrative mediation tends also towards decision-making. Consequently, the choice of participants is particularly important. For instance, the choice of the public authority's representative will often be key to the mediation success; not only will a line manager or an elected representative have the power to negotiate, but they are more likely to have the necessary knowledge and experience for creative problem-solving. In addition, private disputants will normally bring legal representation, which is key in redressing the inherent imbalance of power. Furthermore, the Grenoble pilot revealed the important role played by consultation with other public bodies: either to elicit ideas for viable solutions (in public

45 See the mediation experiences shared by a wide range of judges, barristers, mediators and civil servants during the 'First Symposium on Administrative Mediation' – Conseil d'Etat, Wednesday 18 December 2019 <www.conseil-etat.fr/actualites/actualites/premieres-assises-nationales-de-la-mediation-administrative>.

46 Administrative mediation is not the norm in many administrative justice systems, and, even when it is 'officially' supported therein, it does not always thrive. In England and Wales, mediation may be cited in the pre-action protocol for judicial review, but it is rarely used.

employment disputes, the centre for local government's civil service is often consulted to this effect), or because their decision may hold the key to the resolution process (as happened with the consultation of the *Bâtiments de France* in respect of repairs to a listed building) or to facilitate compliance with the mediation agreement (e.g. endorsement of the financial settlement by the public accountant or finance department).

Finally, in some cases, the participation of 'interested parties' beyond the disputants will be necessary in order to solve the dispute at all. In such case, mediation can become a polycentric or participatory decision-making process, certainly better suited to resolving disputes with multiple interests than the judicial process. Thus, the contact with administrative law may transform mediation into a proteiform and informal decision-making process.

This shows that, rather than mediation being unsuited to administrative law, it can be structured in useful ways to address the challenges of administrative disputes and to adapt to the specificities of public administration.

5.1.2 The mediation agreement

Importantly, the drafting, validity and enforcement of the mediation agreement are also shaped by the public law nature of the dispute. Both the agreement's specific characteristics and its approval process result from the tensions between confidentiality – a key principle of mediation – and accountability – a key principle of public law.

In Grenoble, more often than not, the mediation agreement was submitted for approval to the local government's elected assembly. In such circumstances, the drafting of the mediation agreement can be tricky. Practices vary[47] but, generally, the need for approval by a political authority determines the succinct nature of the 'official' agreement so as to protect a degree of confidentiality. This also explains Article L213-2 of the 2016 law: it states that mediation respects the principle of confidentiality but allows an exception when implementation of the agreement requires its existence and/or content to be divulged.

Approval by a political authority was often sought to strengthen the validity of the mediation agreement. As Grenoble's court (and the national developing practice) discourages the systematic certification of mediation agreements, the formal endorsement by an elected assembly serves to replace the legal validity that such a certification would have granted. The move from legal to political accountability and from legal validity to political legitimacy facilitates enforcement and compliance.[48] The question of validity and legitimacy of the mediation agreement is to be explored in the next section, but the shift analysed above is key to defusing the tensions specific to administrative mediation.

5.2 Resolving the tensions between the mediation and adjudication paradigms

I previously postulated that mediation and adjudication belonged to conflicting paradigms with adjudication typically being hierarchical, top-down and rules-driven and mediation being horizontal, bottom-up and problem-solving.[49] While these

47 Sometimes two versions of the agreement are drafted: one – succinct – for 'political' approval and one for the purpose of the mediation process.

48 This is particularly important in view of the problems of compliance that some court judgments encounter.

49 See Boyron (n 4).

representations of mediation and adjudication are not original,[50] they are particularly significant in the context of administrative law. Not only do they lead to tensions of the type identified in the previous section, but the competition between paradigms has the potential to disrupt the wider administrative justice system.

When considering the (then) new representation of legal systems as networks as opposed to the more classical and pyramidal representation, Ost and van de Kerchove examined the validity of norms in the new 'network' paradigm.[51] They explained that the hierarchical, exclusive and unilateral process of validation in the classical paradigm of a legal system did not match the validation process of the network paradigm. The latter is guided instead by reference to formal validity (i.e. legality), empirical validity (i.e. effectiveness) and meta-validity (i.e. legitimacy), with the validity of each norm being determined by a specific equilibrium between the three criteria, and this equilibrium itself being dynamic and subject to change.[52] What the French experience shows is that these validity criteria and their equilibrium play a significant role in administrative mediation.

Throughout the mediation process, the validity criteria of effectiveness and legitimacy are largely dominant; they underpin the search for a solution, with the criterion of legality playing a lesser role. However, once an agreement is reached, the balance between the three validity criteria threatens to change: the need for enforcement of the agreement tends to give dominance to the criterion of legality. A mediation agreement the validity of which arises primarily from its effectiveness and legitimacy loses this validation when scrutinised against the criterion of strict legality for its enforcement.

The equilibrium between the three validity criteria is particularly affected in France by the availability of certification recognised in the new law.[53] To certify an agreement, administrative courts check the content against the principles of legality and public order, thereby shifting the balance between the three validity criteria and bringing the two adjudication and mediation paradigms into conflict. While a validity of norms giving pre-eminence to the criteria of effectiveness and legitimacy is able to support a creative, consensual and problem-solving solution, the emphasis on formal validity for the certification process reasserts the dominance of legality, thereby threatening all that can be achieved in mediation. The tensions arising from this shift are such that, if not neutralised, mediation may struggle to thrive.

In Grenoble, Vice-President Wegner seemed intuitively aware of these tensions. In practice, once a mediation was agreed to, the dispute would only rarely be the subject of certification and would not normally return to the court, and so would not 're-integrate' the adjudication process (unless and until mediation failed). This policy of 'certification avoidance' managed the tensions arising from certification, with each dispute resolution mechanism (be it mediation or adjudication) remaining on a parallel course. The Grenoble court mandated notification of the existence of an agreement to close the court case, but as outlined in part 3 it did not enquire into the agreement's content. Importantly, the lack of certification had no implications for compliance throughout the pilot.

Finally, the approval of mediation agreements by elected representatives completes this process of neutralisation of the tensions between the two paradigms: the shift from

---

50 See, for instance, C Menkel-Meadow, 'Dispute resolution' in P Cane and H M Kritzer (eds), *The Oxford Handbook of Empirical Legal Research* (Oxford University Press 2010) 596.

51 F Ost and M van de Kerchove, *De la pyramide au réseau? Pour une théorie dialectique du droit* (PU Saint Louis 2002).

52 Ibid 307.

53 See Article L213-4 CAJ.

legal to political accountability and from legal validity to political legitimacy facilitates compliance, counteracts the absence of certification and avoids a shift in the validity criteria, thereby ensuring that the dispute remains with the mediation paradigm. Thus, the experience in France shows that despite adjudication and mediation being paradigmatically opposed, careful management of the moments of trans-paradigm interaction – including minimising the formal adjudicatory process once mediation is underway and creating instead internal validation processes within mediation – can enable both to operate in parallel. However, this should not be understood as sanctioning 'illegal' agreements, but rather as managing paradigmatic tensions that would otherwise have the power to stifle the problem-solving mediation process upstream.

This experience also tells us that what is needed is a 'negotiated environment' within which administrative mediation can operate.

5.3 Administrative mediation: the creation of an 'alternative' system of administrative justice?

The French experience suggests that, to thrive, administrative mediation needs to be embedded in a 'negotiated environment' with dedicated professionals, new networks of stakeholders and the legitimation of a distinct ideology, albeit within the institutional structures of administrative law. This in turn has resulted in the emergence of new dispute resolution practices, understandings and cultures (both political and legal). Thus, the introduction of administrative mediation becomes a major enterprise in 'informal ordering' that resembles the launch of an 'alternative' system of administrative justice.[54] This becomes clear when the constituent parts of this 'alternative system' are analysed.

5.3.1 An ideology of administrative mediation

In France, administrative mediation rests largely on the assumption that administrative courts are ill-suited to some disputes, and in turn that mediation is well suited to them. As we have seen already, this underpinned both the Grenoble pilot and the national implementation. However, what appears at first to be a rationale for introducing administrative mediation should be understood as an ideology that frames and legitimises the existence of mediation in the administrative justice system. Harrington and Merry[55] demonstrated that systems of mediation are not only ideologically constructed but that they produce ideology in and of themselves. Similarly, the Grenoble pilot began the production of a discourse and supportive ideology to explain, legitimise and circumscribe the resort to administrative mediation, an ideology which is both distinct from and fitting with the one underpinning the role of French administrative courts. By focusing on disputes that are not suited to the administrative court system, this ideology provides both a clear delineation between (sub-)systems of administrative justice and substantive legitimation; mediation's very existence is not only sanctioned but required by the existing system of administrative justice. Furthermore, this choice provides a narrow ambit for the use of mediation: the majority of cases in the administrative courts will not be seen as suitable for mediation. Mediation is recognised as well as corralled by this choice of underpinning: adjudication remains unassailably mainstream.

54 While administrative justice has two contrasting meanings – see S Halliday and C Scott, 'Administrative Justice' in Cane and Kritzer (n 50) – in the present article, administrative justice refers to the various methods of redress against administrative decisions (e.g. courts, ombudsman, complaint procedures, mediation); beside the mechanisms of dispute resolution, this system encompasses dedicated professionals and stakeholders, an ideology as well as pathways, practices, culture and understandings.

55 See C Harrington and S Merry, 'Ideological production: the making of community mediation' (1988) 22 Law and Society Review 709.

Interestingly, this ideology is completed by the more recent case law of the Conseil d'Etat. Since 2017, the Conseil d'Etat has repeatedly stated that there exists: '… a mission of public interest pertaining to the State to develop ADR mechanisms, a corollary to a good administration of justice'.[56] This provides a clear endorsement of administrative mediation by the Supreme Court and pertains to the construction of this ideology: mediation has intrinsic value due to its public interest mission, and it is placed at the core of the French justice system. It receives the full backing of the state, thereby creating a complex system of legitimacy and ideology.

5.3.2 The new professionals and stakeholders of administrative mediation

The Grenoble pilot revealed the key role played by dedicated professionals in the success of mediation. It is not surprising therefore that the introduction of administrative mediation resulted in the emergence of new categories of professionals: mediation champions in the administrative courts, as well as specialist 'administrative mediators' (i.e. persons who are both experienced mediators and experts in the administrative law field in dispute and thus qualified under the new law). Not only is a national list of mediators being established, as seen in part 3, but mediation providers, such as the *Chambre nationale des praticiens de la mediation*, have started to offer the services of 'administrative mediators'.[57]

Also, the introduction of mediation has required administrative courts to work with new stakeholders or work differently with established ones. For instance, through the 2013 Agreement, Grenoble has built a network of stakeholders (old and new) to provide both structure and impetus for the system's acclimatisation to mediation. Since then, the court's local strategy has extended its network(s) by negotiating new agreements and by establishing new contacts (especially with public authorities). As all 42 administrative courts have been encouraged to follow Grenoble's lead, negotiate their own local agreements, deploy their networks and in some cases even help the emergence of local mediation providers, this phenomenon is spreading through France.

5.3.3 Working practices and culture

In Grenoble, administrative mediation has resulted in the creation of new pathways inside the court and changes to the working practices of many stakeholders. As seen above, the practice by local governments to approve the mediation agreements has created new mechanisms of accountability and has started to change local political culture.

The success of administrative mediation relies also on transforming the very culture of dispute resolution. There are early signs of this in the legal professions. This was somewhat apparent in Grenoble, with several public law barristers reporting an increased interest in solving the dispute (rather than winning the case). In fact, several mediations were triggered at the request of barristers. This also reflects a wider involvement of the French legal professions with mediation as successive legislative reforms have made this form of dispute resolution increasingly mainstream. With more local Bars offering mediation provisions, several have signed an agreement with the local administrative court to become the mediation provider.

56 See, for instance, CE 17 March 2017 no 403768 and CE 21 October 2019 no 430062.

57 For instance, on its website, the Chambre nationale des praticiens de la médiation lists its administrative mediators separately with their training and area of expertise: see <www.cnpm-mediation.org>.

5.4 A plural administrative justice system

This suggests that administrative mediation is emerging as a distinct (sub)system, operating alongside the administrative courts. In France, administrative mediation is on the way to creating a nationwide informal ordering of dispute resolution and, thus, is introducing a degree of plurality in the administrative justice system. This may mark the beginning of a wider recalibration of administrative justice there.

5.4.1 Administrative mediation as state law pluralism

As mediation has consistently been associated with legal pluralism[58] and informal legal ordering,[59] I believe the Grenoble experience exhibits features of legal pluralism as explained by Gad Barzilai:

> [Legal pluralism] primarily articulates detachment from legal centralism revolving around state law, criticism of the exclusiveness of state law, decentralization of court-centered judicial studies, exploration of non-state legal orders, unveiling of informal socio-legal practices, and an understanding of law as a multi-centered field that deals with the convergence of a multiplicity of norms, localities, states, global sites and practices.[60]

The proclaimed aim of the pilot (and of the 2016 reform) was to introduce an alternative mechanism for the resolution of those administrative law disputes that are not well catered for in the courts and to recognise a key role for disputants in developing solutions outside the judicial process; as such these indicate a decentralisation of dispute resolution away from the courts and hierarchical judicial process, a key feature of legal pluralism.

Furthermore, as seen above, the Grenoble pilot has generated any number of informal socio-legal practices giving rise to an ordering that did not exist previously. In effect, this initiative triggered a parallel system that has all the hallmarks of an alternative if local system of administrative justice. Furthermore, having informed substantially the implementation of the 2016 reform in an experimental and bottom-up approach, early indications show this informal system spreading through France. Consequently, there is a strong case for recognising as instances of legal pluralism, the Grenoble pilot and arguably, the introduction of administrative mediation throughout France. In fact, this new informal ordering at the heart of the state's administrative justice system would seem to pertain to state law pluralism, i.e. the recognition of plurality *within* (and not simply alongside) a state's legal order.

At this stage, it is important to highlight that not only is the literature on legal pluralism deeply conflicted,[61] but that some proponents of legal pluralism have rejected the possibility of distinct legal orderings within a state's legal order and of plurality within state law, thereby rejecting the notion of state law pluralism.[62]

---

58 See Harrington and Merry (n 55) 709.

59 See D Smith, H Blagg and N Derricourt, 'Mediation in the shadow of the law: the south Yorkshire experience' in R Matthews (ed), *Informal Justice?* (Sage 1988) 123; and C Harrington, 'Informalism as a form of legal ordering' in K Wittington, D Kelemen and G Caldeira (eds), *The Oxford Handbook of Law and Politics* (Oxford University Press 2008) chapter 21.

60 Gad Barzilai, 'Beyond relativism: where is legal power in legal pluralism?' (2008) 9(2) Theoretical Inquiries in Law 395.

61 See, for instance, B Tamanaha, 'The folly of the social scientific concept of legal pluralism' (1993) 20 Journal of Law and Society 192.

62 See G Woodman, 'Ideological combat and social observation – recent debate about legal pluralism' (1988) 42 Journal of Legal Pluralism and Unofficial Law 21; and H B A Sani, 'State law and legal pluralism: towards an appraisal' (2020) 52 Journal of Legal Pluralism and Unofficial Law 82.

However, an analysis deriving from state law pluralism speaks to the growing experience of administrative mediation in France; it is key to understanding the two systems of administrative justice (formal and informal) and their interactions. From this, it may be possible to posit two 'interim' interpretations of these complex dynamics from the perspective of state law pluralism.

The first interpretation sees in administrative mediation an informal (sub)-system of administrative justice first tolerated and then sponsored by the state to expand the capacity and reach of the administrative justice system as a whole. The emergence of a complex and distinct ordering with personnel, pathways, ideology and a distinct accountability process goes far beyond an alternative mechanism of dispute resolution. While this portrayal highlights the relative dependence (and fragility) of the informal ordering, it also hints at its possible autonomous entrenchment. With the introduction of administrative mediation in 42 separate courts, the pluralist potential of the 2016 reform is unleashed further. As networks and practices are likely to diverge from court to court, the emergence of local alternate systems will lead to more plurality.

The second interpretation treats the 2016 reform with suspicion, highlighting the fact that the Grenoble experiment has now been captured, transformed and re-imposed in a top-down manner via state legislation, thus losing its plural and disruptive character.[63] Furthermore, the support and energy deployed by all administrative courts (particularly the Conseil d'Etat) to ensure the success of administrative mediation undermines the claim that it is an enclave of state law pluralism. The same goes for the relation that administrative mediation entertains with case law: the labelling of mediation's development as a public interest mission may be interpreted as the beginning of a reclaiming process through law, thereby questioning the pluralist aspirations of this national deployment of administrative mediation.

The future will tell which of these two interpretations is accurate and whether the plurality of the French administrative justice system is a reality.

5.4.2 Remodelling informal administrative justice?

Not only has administrative mediation introduced plurality into the French system of administrative justice, but its unusual characteristics could be instrumental in shaping further reflection on informal administrative justice.

Administrative mediation is commonly (and rightly) categorised as an alternative dispute resolution mechanism alongside ombudsmen, arbitration etc. In this context, informal administrative justice has been the subject of several studies, some concentrating on one specific mechanism and others spanning the wider range of ADR.[64] Administrative mediation would certainly benefit from better conceptualisation as part of informal administrative justice and from recent studies in the field. For instance, the use of the restorative justice framework may prove pertinent to analysing and conceptualising French administrative mediation and its practice.[65]

In addition, I wish to argue that administrative mediation is unusual in that it straddles the divide between remedies and decision-making that commonly drives the

63 See, for instance, the critical stance taken by Richard Abel in the introduction of his edited volume, The *Politics of Informal Justice* vol 1 – The American Experience (Academic Press 1982).

64 See Naomi Creutzfeldt, *Ombudsmen and ADR: A Comparative Study of Informal Justice in Europe* (Palgrave 2018).

65 See Naomi Creutzfeldt, 'A voice for change: trust relationship between ombudsmen, individuals and service providers' (2016) 38 Journal of Social Welfare and Family Law 460.

conceptualisation's effort of administrative justice.[66] As already shown above, the Grenoble pilot shows that administrative mediation goes beyond providing redress for a past action or decision and aims to find a permanent solution to disputes into the future. To this effect, administrative mediation is both a remedial tool and a flexible decision-making process. This dual character is key to understanding French administrative mediation and needs to drive its conceptualisation.

The literature on administrative justice contains several taxonomies identifying the different conceptions of administrative justice that drive types of decision-making and some can be applied appropriately to administrative mediation.[67] For instance, Kagan's study[68] of the role of (mis)trust in the design and operation of decision-making informs a typology where administrative mediation would fit well the ideal-type of 'negotiation/mediation' that combines legal informality with participation by individuals or organisations in decision-making. It is clear that this typology speaks to the French experience. Furthermore, the concept of trust, which is key to his analysis of agencies' decision-making styles places legal creativity at the nexus of strong adherence to rules and emphasis on social values: this would certainly provide a grid of explanation for the search for bespoke/creative solutions within the existing legal framework. Moreover, this work has the added advantage of giving a central role to the concept of trust, a concept that could be particularly pertinent in conceptualising both the remedial and decision-making aspects of administrative mediation. Indeed, it may be that conceptualisation of French administrative mediation could play a role in helping us rethink administrative justice beyond the remedy/decision-making divide.

Conclusion

After years of false starts, administrative mediation seems to be finally taking root in France with the 2016 reform. The Grenoble pilot that informed its implementation shows that mediation can work in the administrative context and that it can fill a gap in French administrative justice. However, the French experience also indicates that to thrive administrative mediation needs to be supported by its own system of administrative justice, thereby leading to a plural administrative justice system in France. Should this distinct (sub)system of informal administrative justice take root nationwide, it may be instrumental in rethinking more widely our modelling of informal administrative justice.

66 See Halliday and Scott (n 54).

67 See, for instance, J L Mashaw, *Bureaucratic Justice: Managing Social Security Disability Claims* (Yale University Press 1983); M Adler, 'A socio-legal approach to administrative justice' (2003) 25 (4) Law and Policy 323; S Halliday and C Scott, 'A cultural analysis of administrative justice' in M Adler (ed), *Administrative Justice in Context* (Hart 2010).

68 R Kagan, 'Varieties of bureaucratic justice' in N Parillo (ed), *Administrative Law from Inside Out* (Cambridge University Press 2017).

NILQ 71(3): 481–499

The Singapore Convention: a solution in search of a problem?

BRYAN CLARK

University of Newcastle, UK

AND

TANIA SOURDIN*

University of Newcastle, Australia

Abstract

This article explores the purpose and efficacy of the United Nations Convention on International Settlement Agreements Resulting from Mediation (the Singapore Convention or Convention). The Convention's genesis was premised on the notion of alleviating the enforceability issues that are annexed to international mediation settlement agreements (IMSAs) arising from cross-border mediation. While such enforceability issues are not entirely unfounded, the way in which the Convention has been drafted to address such issues has been the subject of criticism. In view of such criticism, this article explores the empirical research upon which the Convention's introduction is based and queries whether the structure of the instrument heralds an unnecessary juridification of the mediation process. In particular, a close review of the research highlights the unintended consequences that can flow from the Convention's uptake, suggesting that the introduction of the Convention may lead to an increase in issues pertaining to IMSA enforcement. It is in this context in which this article submits that the Convention may be regarded as a solution in search of a problem.

Key words: mediation; alternative dispute resolution (ADR); cross-border mediation; Singapore Convention; international commercial mediation; barriers to mediation; enforcement.

Introduction

This article is concerned with the UN Convention on International Settlement Agreements Resulting from Mediation (the Singapore Convention or Convention), a multinational treaty with its origins in the UN Commission on International Trade Law (UNCITRAL). In short, the Singapore Convention introduces a new, uniform mechanism for the cross-border enforcement and recognition of settlement agreements rendered in international commercial mediations – international mediation settlement agreements (IMSAs). In this sense, the Singapore Convention broadly mirrors and takes inspiration from the approach found in the New York Convention for the cross-border enforcement of arbitral awards in international commercial arbitration.[1] Some four years in gestation, the

* Bryan Clark email: bryan.clark@newcastle.ac.uk. Tania Sordin ORCID number <https://orcid.org/0000-0001-8620-1513>. The authors would like to thank the anonymous referees for their comments on an earlier draft of this article.

1 UNCITRAL, 'Convention on the Recognition and Enforcement of Foreign Arbitral Awards' (10 June 1958) 330 UNTS 38.

Singapore Convention was adopted by the UN General Assembly on 20 December 2018[2] and signed by 46 countries[3] in Singapore on 1 August 2019, although at the time of writing the number of signatories has climbed to 53. The Convention came into force on 20 September 2020, which occurred six months after deposit of the third instrument of ratification, acceptance, approval or accession.[4] Concurrently, amendments were also made to the Model Law on International Commercial Conciliation (the Model Law) which can be adopted by states to help ease the Convention into their domestic systems.[5] The Singapore Convention has been broadly welcomed by mediation organisations and commentators alike as providing a useful 'shot in the arm' for international commercial mediation.[6]

Despite this enthusiasm, the terms of the Convention's drafting have not escaped academic and professional criticism.[7] While in this article some facets of this critique are explored, the central task is to scrutinise the need for the Convention in the first place and to examine some of the potentially deleterious unintended consequences arising from its introduction. In short, the authors posit that the Convention could be regarded as a solution in search of a problem. It is thus argued that the premise behind the Convention's introduction is based both upon questionable empirical evidence as well as potential misapprehension of the nature of mediation by the uninitiated. Furthermore, while accepting that the Singapore Convention may provide greater publicity and enhanced legitimacy to international commercial mediation through its regulatory functions,[8] the authors caution that this new instrument may herald an unnecessary juridification of the process and lawyer domination therein that may paradoxically lead to *greater* enforcement problems with mediated outcomes and militate against some key, qualitative benefits of the process. Other barriers to uptake are also explored, as well as measures to help expedite acceptance and use of mediation in international commercial matters, such as awareness-raising, overcoming cultural and professional barriers and greater embedding within contracts, that should not be forgotten in the rush to embrace the Singapore Convention.

The structure of this article is as follows: first, the origins and nature of the Singapore Convention are mapped out with a brief examination of its main provisions; this is followed by an analysis of the evidence base behind the need for a uniform enforcement

---

2 See UN General Assembly, 'Report of the United Nations Commission on International Trade Law' (25 June–13 July 2018) UN Doc A/73/17, 50–55.

3 Including the USA and China. See, generally, Penningtons Manches Cooper LLP, 'The Singapore Convention on Mediation: could 2020 be the year of the ratification?' (*Lexology*, 22 January 2020) <https://www.lexology.com/library/detail.aspx?g=43578b07-7b21-4ba7-a229-a7f98ed0254d>.

4 UNCITRAL, 'United Nations Convention on International Settlement Agreements Resulting from Mediation' (20 December 2018) UN Doc A/Res73/198, Article 14(1). Six states have ratified the Convention at the time of writing.

5 UN, *UNCITRAL Model Law on International Commercial Conciliation and Guide to Enactment and Use 2002* (United Nations 2004) 36–37, paragraph 55.

6 See, for example, Eunice Chua, 'The Singapore Convention on Mediation – a brighter future for Asian dispute resolution' (2019) 9 Asian Journal of international Law 195; Gary Birnberg, 'Singapore Convention brings big changes for litigators and arbitrators' (*JAMS*, 5 August 2019) <www.jamsadr.com/blog/2019/singapore-convention-brings-big-changes-for-litigators-and-arbitrators>; Haris Meidanis, 'International enforcement of mediated settlements: two and a half models – why and how to enforce internationally mediated settlement agreements' (2019) 85(1) Arbitration (London) 49.

7 See, for example, Migle Zukauskaite, 'Enforcement of mediated settlement agreements' (2019) 111 Teise 205, 211–214; Masood Ahmed, 'Reflections on the UNCITRAL Convention on the Enforcement of Mediation Settlement Agreements and Model Law' (2019) Lloyds Maritime and Commercial Law Quarterly 259.

8 Indeed, this is seen as especially important in some jurisdictions such as Asia: see generally, Gloria Lim, 'International commercial mediation: the Singapore model' (2019) 31 Singapore Academy of Law Journal 377.

mechanism for IMSAs; and finally an examination is offered of some of the potential deleterious consequences arising from introduction of the Singapore Convention, with a brief discussion of other barriers to uptake and strategies that may help expedite the greater use of international cross-border mediation.

1 Provisions of the Singapore Convention

1.1 Basic premise of the Convention

The terms of the Convention were the subject of extensive negotiation and revision over its period of gestation. In this sense, the Convention is inevitably a product of compromise.[9] Some of the consequences that this hard-fought bargaining had on the drafting of the Convention are discussed further below. It is first, however, worth noting that the Convention can be invoked as both a 'sword' and a 'shield'. In the former sense, it may be used as a way in which to enforce mediated outcomes and, in the latter sense, as a means of resisting legal action by another on the basis that the dispute has already been resolved. In this regard, Article 3(1) allows parties to instigate proceedings in signatory states to enforce the terms of the relevant IMSA. Article 3(2), by contrast, allows parties to invoke the terms of an IMSA in order to prove that a matter by which another party is seeking a remedy has already been resolved.

1.2 Definitional Issues

Article 1 provides that the Convention applies 'to an agreement resulting from mediation and concluded in writing by parties to resolve a commercial dispute ("settlement agreement") which, at the time of its conclusion, is international'. Article 1 goes on to tie the definition of 'international' to the settlement agreement (rather than the subject of the dispute). It provides that the agreement is international in that:

> (a) At least two parties to the settlement agreement have their places of business in different States; or (b) The State in which the parties to the settlement agreement have their places of business is different from either (i) The State in which a substantial part of the obligations under the settlement agreement is performed; or (ii) The State which the subject matter of the settlement agreement is most closely connected.

There is no attempt in the Convention to adopt the notion of a particular mediation 'seat'. The Convention provides for a form of direct enforcement in that a party to an IMSA can seek enforcement in any country it chooses with no need for a process of review in the country of origin.[10] As has been noted, this model 'essentially delocalises from the enforcement process the place where the [IMSA] may have been reached … This has the extra value that it can be of use to the existing and increasing electronic mediation proceedings that parties in mediation expect to have, so as to design solutions not tied to a specific legal system.'[11]

While the term 'commercial' is not defined in the Convention, reference can be made to the Model Law where the term is afforded a wide interpretation.[12] Certain kinds of disputes are specifically excluded from the scope of the Convention. According to Article 1(2), the Convention does not apply to settlement agreements '(a) Concluded to

9 See Timothy Schnabel, 'The Singapore Convention on Mediation: a framework for the cross-border recognition and enforcement of mediated settlements' (2019) 19 Pepperdine Dispute Resolution Law Journal 1.

10 Meidanis (n 6) 53.

11 Ibid 53.

12 Model Law, Article 1.

resolve a dispute arising from transactions engaged in by one of the parties … for personal, family or household purposes; (b) Relating to family, inheritance or employment law.' Furthermore, Article 1(3)(a) provides that the Convention shall not apply to settlement agreements: '(i) That have been approved by a court or concluded in the course of proceedings before a court; and (ii) That are enforceable as a judgement in the State of the court.' Similarly, Article 1(3)(b) excludes 'Settlement agreements that have been recorded and enforceable as an arbitral award.' The narrowing of the Convention's reach in these regards has attracted some adverse commentary. For example, some commentators have noted that the exclusion of consumer and employment disputes is an opportunity lost to help expedite growth in these areas and that the carve-outs, particularly around IMSAs otherwise enforceable in court regimes, are manifestly uncertain.[13] This may in fact mean that some IMSAs fall through the gaps as not enforceable under the Singapore Convention but in practice will not be enforceable through the Hague Convention either.[14]

The term 'mediation' is given a wide definition in the Singapore Convention to encapsulate a range of processes, irrespective of the domestic term used to describe them, in which 'parties attempt to reach an amicable settlement of their dispute with the assistance of a third person … lacking the authority to impose a solution upon the parties'.[15] Such flexibility is also found in the requirement that the IMSA be recorded in writing. Under Article 2(2), this is defined as being recorded in 'any form', specifically encompassing an agreement found in 'electronic communication' such as email correspondence.

1.3 Proof of IMSA

A party seeking to rely on an IMSA must evidence its existence. Thus, the agreement must be signed by the parties.[16] Evidence must also be provided that any agreement stemmed from mediation. This can be achieved in accordance with Article 4(1)(b) by either:

> (i) the mediator's signature on the settlement agreement; (ii) a document signed by the mediator indicating that the mediation was carried out; (iii) an attestation by the institution that administered the mediation; or (iv) in the absence of (i), (ii) or (iii), any other evidence acceptable to the competent authority.

Article 4(2) makes special provision meeting the requirements for signature in respect of electronic communication.

1.4 Grounds of relief

A range of grounds under which an IMSA may be deemed unenforceable are included in the Convention under Article 5. It has been noted that, while such exceptions should provide succour to states to ensure that settlements can be challenged in appropriate circumstances, the wide range of potential caveats may lead to significant post-mediation disputes regarding enforcement.[17] As discussed later in this article, one of the dangers here is that such provisions may in fact concentrate the mind of parties and their lawyers on the possibility of post-settlement challenges to enforcement. Clearly, such consequences are contrary to the aims of the Convention.

13 See, generally, for example, Zukauskaite (n 7).

14 The Hague Convention of 30 June 2005 on Choice of Court Agreements (concluded 30 June 2005) HCCH (The Hague). For a consideration of some of these issues, see, for example, Zukauskaite (n 7) 212.

15 UNCITRAL (n 4) Article 2(3).

16 UNCITRAL (n 4) Article 4(1)(a).

17 Zukauskaite (n 7) 213; Ahmed (n 7) 267–269.

It is true that many of these grounds represent familiar territory and replicate those found in the New York Arbitration Convention. By dint of Article 5(1), these include: incapacity; when the agreement is null and void, inoperative, or incapable of being performed; that the agreement is not binding or not final according to its terms; that it has been subsequently modified; that the obligations have been performed or are not clear or comprehensible; and that relief would be contrary to the terms of the settlement agreement. Under Article 5(2), exceptions are also provided for where granting relief would be contrary to public policy, or the subject matter is not capable of settlement by mediation of the competent authority.

Further grounds for refusing to grant relief pertain to circumstances in which there has been either a 'serious breach of mediator standards'[18] or 'a failure by the mediator to disclose to parties circumstances that raise justifiable doubts as to the mediator's impartiality or independence'.[19] These provisions have received a mixed reception. As Meidanis has noted, these Articles 'have been criticised for dangerously extending the control of the MSA by the state of enforcement'.[20] One commentator has suggested that these grounds 'can be interpreted as a map for counsel to relieve its client of its obligations of a settlement agreement by focusing on the conduct of the mediator' and cautioned that attempts to unpick settlements on such bases may erode mediation's important principles of confidentiality and evidential privilege.[21] Others have played down such concerns on the basis that the use of these measures is limited to occasions in which parties must establish that, without the alleged serious breach of standards or failure to disclose circumstances laying doubt on impartiality or independence, the party concerned would not have entered into the mediation agreement.[22] It is certainly the case that these provisions will raise questions around requirements for applicable mediator standards.

1.5 Opt out/opt in

A further controversial provision is Article 8 and, in particular, the fact that signatories may stipulate that the Convention shall apply 'only to the extent that the parties to the settlement agreement have agreed to the application of the Convention'.[23] While this provision was required to accommodate the diverse views of delegates to the UNCITRAL Singapore Convention discussions, it has been criticised by some commentators.[24] It has been argued that operation of this provision may lead to an uneven application of the Convention and also afford the opportunity to more dominant parties to coerce their counterparty to opt in to the terms of the Convention or otherwise to ensure that they can follow the most advantageous approach to pursuing or resisting enforcement of the IMSA.[25]

18 UNCITRAL (n 4) Article 5(e).

19 Ibid Article 5(f).

20 Meidanis (n 6) 56.

21 F Peter Phillips, 'Concerns on the new Singapore Convention' (*Mediate*, October 2008) <www.mediate.com/articles/phillips-concerns-singapore.cfm>.

22 See, for example, Karl Mackie, 'Another historic step for mediation' (*LinkedIn Pulse*, 8 August 2019) <www.linkedin.com/pulse/another-historic-step-mediation-karl-mackie-cbe> accessed 15 June 2020.

23 UNCITRAL (n 4) Article 8(1)(b).

24 Schnabel (n 9) 56.

25 See Jan O'Neil, 'The new Singapore Convention: will it be the New York Convention for Mediation?' (*Thomson Reuters, Practical Law: Dispute Resolution Blog*, 19 November 2018) <http://disputeresolutionblog.practicallaw.com/the-new-singapore-convention-will-it-be-the-new-york-convention-for-mediation>.

2 Why is the Convention needed?

2.1 The perceived need

The basis for the Singapore Convention's introduction is, as we will articulate below, well established. It can be argued, however, that principally the rationale emanates from a perception of failure. This failure, in particular, relates to the notion that commercial mediation has not made significant strides in the cross-border context.[26] This may seem particularly galling to mediation enthusiasts when set against the relatively strong growth of domestic commercial mediation found across a range of jurisdictions.[27] In terms of explaining this disappointing state of affairs, for some time mediation providers and commentators in the field have expressed the idea that a major stifling factor to development of the process in the cross-border context has been the lack of any uniform enforcement mechanism for mediated agreements akin to that which operates in the field of international arbitration through the New York Convention.[28] Indeed, views in favour of the desirability of a new uniform enforcement mechanism were aplenty in the inter-party UNCITRAL negotiations leading up to the adoption of the Singapore Convention.[29]

Mediation outcomes have historically been subject to quite varying treatment across different jurisdictions in terms of their enforcement. A range of possibilities exist including, most fundamentally, enforcement as contracts. Mere enforcement under general contract law brings with it, however, the uncertainties that can arise around conditions for enforcement and the general grounds for challenge such as fraud, misrepresentation, error and duress.[30] As an improvement on this general situation, in some jurisdictions, outcomes rendered in mediation may be conferred a special enforcement status by the use of settlement deeds and notarised arrangements. In others, mediated settlements may by expressly approved by court. However, the situation is highly variable across different jurisdictions.[31] As Zukauskaite notes, within the EU 'every country has a mechanism how to transform a [a mediated settlement agreement] into a directly enforceable title'.[32] However, '[o]utside the EU situation varies significantly'[33] with some countries having no means of transferring a mediated settlement into a directly

26 See, for example, International Institute for Conflict Prevention and Resolution and Centre for Effective Dispute Resolution, 'Insights into Alternative Dispute Resolution' (Winter 2018–2019) 1, 3 <www.cpradr.org/news-publications/reports/2019-04-04-cpr-cedr-joint-insights-report-on-the-use-of-adr/_res/id=Attachments/index=0/CEDR%20CPR%20report%20040419.pdf>; Kim Shi Yin, 'From "face-saving" to "cost saving": encouraging and promoting business mediation in Asia' [2014] 32(10) Alternatives to the High Cost of Litigation 158, 158.

27 For a review of developments, see Neil Andrews, *The Three Paths of Justice: Court Proceedings, Arbitration, and Mediation in England* (2nd edn, Springer 2018).

28 David Weiss and Brian Hodgkinson, 'Adoptive arbitration: an alternative approach to enforcing cross-border mediation settlement agreements' (2014) 25 American Review of International Arbitration 275; Lynn Cole, 'Exploring international mediation' in Alexia Georgakopoulos (ed), *The Mediation Handbook: Research, Theory, and Practice* (Routledge 2017) 318–319.

29 For a review of these discussions, see Schnabel (n 9) 3–4.

30 For a discussion in the US context, see Edna Sussman, 'The final step: issues in enforcing the mediation settlement agreement' (2008) The Fordham Papers 1, 7–16 <https://sussmanadr.com/docs/Enforcement_Fordham_82008.pdf>.

31 For a review of different approaches, see Bobette Wolski, 'Enforcing mediated settlement agreements: critical questions and directions for future research' (2014) 7(1) Contemporary Asia Arbitration Journal 87, 93–99.

32 Zukauskaite (n 7) 207.

33 Ibid 207.

enforceable title.[34] While such measures may be generally considered adequate in respect of the enforcement of domestic mediated outcomes, the situation is more complex with regards to cross-border mediated agreements, with some of the delegates to the Convention negotiations highlighting such problems.[35]

In light of the perceived problem in terms of enforceability of IMSAs, creative initiatives have been put in place to render mediated outcomes enforceable as 'consent awards' under the New York Convention through 'arb-med-arb' proceedings.[36] Such approaches entail a dispute first brought before an arbitral tribunal, with the arbitration then stayed immediately for mediation, and if a solution is brokered the matter is transferred back to the arbitration tribunal for recording as a consent award. Although there are varying views expressed as to whether consent awards are generally enforceable under the New York Convention for want of a dispute at the time the consent award is put in place by an arbitrator,[37] it has been suggested that beginning the process in arbitration is sufficient to overcome these concerns by meeting the terms of Article 1 of the New York Convention as applying to matters 'arising out of differences between persons'.[38] Even if such measures may be successful in a technical sense, the blending of mediation with arbitration in this way can lead to increased party costs and process complications which may limit their effectiveness.[39] The potential harm arising from the infusing of mediation with more legal processes is a matter that we develop below.[40]

2.2 Empirical studies

A range of studies were undertaken in the lead-up to the adoption of the Singapore Convention which, at first glance, lend succour to the notion that a unified international enforcement mechanism would be beneficial for future growth of international commercial mediation.

One of the principal studies that was conducted specifically to aid UNCITRAL's consideration of the new enforcement mechanism was that undertaken by Strong.[41] Strong's work analysed survey responses from some 221 respondents drawn from a range of jurisdictions, including the USA, the UK, across Europe, Asia, Latin America, Oceania and the Middle East.[42] Respondents included private practitioners, third-party neutrals, academics, internal counsel and those in 'other forms of employment', including judges, government officials, those employed by arbitration institutions.[43] Just over half of the respondents had been involved in at least one international commercial mediation over

34 Ibid 208.

35 See Schnabel (n 9) 3.

36 Including under the Singapore International Mediation Centre Protocol 2014. See for a discussion on the 'arb-med-arb' clause: 'The Singapore International Mediation Centre' (*Singapore International Mediation Centre Insights*, 2 November 2014) <http://simc.com.sg/blog/2014/11/02/singapore-international-mediation-centre>.

37 See the discussion in Sussman (n 30) 21–23.

38 Zukauskaite (n 7) 209–210.

39 See, for example, Bobette Wolski, 'Arb-med-arb (and MSAs): a whole which is less than, not greater than the sum of its parts?' (2013) 6(2) Contemporary Asia Arbitration Journal 249.

40 See section 3.3 below.

41 S I Strong, 'Realizing rationality: an empirical assessment of international commercial mediation' [2016] 73 Washington and Lee Law Review 1973.

42 Ibid 2017.

43 Ibid.

the three years preceding their completion of the survey, with the remainder having had no such engagement in the previous three years.[44]

The survey asked respondents[45] why parties might decline the use of international commercial mediation. Although concerns over the lack of enforceability of IMSAs was one of the issues listed in the survey, this was not included in the top five overall responses for either parties who had previous experience in international commercial mediation or those that did not. Strong explained:[46]

> [T]hose respondents who had been involved with at least one international commercial mediation within the previous three years chose the parties' lack of experience with mediation as the most likely reason why parties did not use that process in the international commercial context, with counsel's lack of experience with mediation coming in as the second most highly ranked option. This group of survey respondents also stated that concerns about revealing litigation or arbitration strategy was the third most likely reason why parties would avoid mediation in international commercial disputes, with concerns about finding an effective mediator coming in as the fourth most popular option.

For those without previous international commercial mediation experience, the most popular responses were a cultural preference for litigation or arbitration.[47] Beyond that it is not clear from the survey data what those without experience in international commercial mediation believe are important factors as to why parties do not use the process.[48] Overall, the data here is far from conclusive as to why parties decline the use of international commercial mediation. Additional questions were asked of those both with and without previous experience about what might encourage the further use of international commercial mediation. For both those with experience and those without, the top three responses were having more evidence of the effectiveness of mediation, better information about the conduct of mediation and better information about the costs of mediation.[49]

Strong's survey also asked respondents directly about the desirability of an international convention on enforcement of IMSAs. Some 74 per cent of respondents viewed that such a development would encourage parties in their own jurisdiction to make use of international commercial mediation and many of the free text comments are supportive of the potential benefits of such an instrument.[50] It is interesting that this idea only really emerged, however, when prompted by a direct question. Nevertheless, the idea that a uniform enforcement mechanism for IMSAs would amount to a boon for international commercial mediation is found in other empirical research. A study conducted by Weiss and Griffiths for the International Mediation Institute (IMI) on the desirability of a uniform enforcement mechanism for IMSAs[51] analysed the views of 103

44 Ibid 2020. Although this was defined as 'including any matter for which the respondent had prepared, even if the mediation was cancelled before actual proceedings began', so it is possible that a proportion of this group did not in fact have actual experience of mediation.

45 This question was directed to all participants, regardless of whether they had participated in a commercial mediation before or not: Strong (n 41) 2033.

46 Ibid 2034.

47 Ibid.

48 Ibid 2035.

49 Ibid 2037–2039.

50 Ibid 2055.

51 David Weiss and Michael Griffiths, 'Report on International Mediation and Enforcement Mechanisms' (Institute for Dispute Resolution IDR (NJCU) School of Business, UNCITRAL Working Group II (Dispute Settlement) 2017).

'users' of commercial mediation drawn from a range of different jurisdictions.[52] Respondents were asked why they believed parties did not attempt to resolve their cross-border disputes through mediation. Of the four options which participants were asked to rank ('they are unfamiliar with mediation'; 'they had a bad experience previously with mediation'; 'they had a bad experience previously with arbitration'; and 'there is no universal mechanism to enforce a mediated settlement'), the most popular response was that of unfamiliarity and the second was the lack of an enforcement mechanism.[53] It cannot be said that these responses are especially informative, however, given the very limited range of options that was offered. Moreover, the option relating to negative past experiences with arbitration seems odd in this context and was unlikely to be popular.

However, when more direct questions regarding enforcement mechanisms were asked, uniformity in participants' responses began to emerge. In response to the question 'Would you be more likely to include a [contractual] mediation clause, if there were a uniform global mechanism to enforce mediation settlements i.e. similar to the 1958 Convention on the Recognition and Enforcement of Foreign Arbitral Awards ("New York Convention")?', 80 per cent of respondents answered in the affirmative.[54] Similarly, in response to the question, 'Would you be more likely to use or increase your use of mediation in a cross border dispute with another party or multiple parties of different jurisdictions if a uniform global mechanism was in place similar to the New York Convention to enforce a settlement agreement reached in the mediation process?', 84 per cent of respondents indicated 'more likely'.[55] The authors also point to similar findings in the Global Pound Conference Series research.[56] In that study, the majority of participants indicated from a set number of options that 'legislation or conventions that promote recognition and enforcement of settlements, including those reached in mediation' was the area 'most likely to improve commercial dispute resolution'.[57]

To what extent such responses were given by potential users, legal advisors or others involved in the field, however, is not apparent from the Global Pound Conference report.[58] Moreover, when the data is broken down by region, some stark differences in views regarding the utility of such legislative intervention is apparent. While legislation promoting the recognition and enforcement of mediated settlements receives a 64 per cent rating from Asia and 61 per cent from Latin America, the percentage drops significantly for other regions such as Oceania (37%) and the UK (41%).[59]

52 The term 'user' is not defined except that it is said to represent 'various field[s] and profession[s] such as law, construction, energy, architecture, international business, healthcare, food and beverage service, water and waste management, tourism, trading, education, and finance': ibid 7, 9.

53 Ibid 11–12.

54 Ibid 14.

55 Ibid 16.

56 Herbert Smith Freehills and PWC, 'Global Pound Conference Series: Global Trends and Regional Differences' (Global Pound Conference Series 2018).

57 Weiss and Griffiths (n 51) 23; Herbert Smith Freehills and PWC (n 56) 14 where the second top answer was 'protocols promoting non-adjudicative processes' (47%) followed by 'cost sanctions' (36%).

58 The five stakeholder groups in the study were '1) Parties: end-users of dispute resolution, generally in-house counsel and business executives; 2) Advisors: private practice lawyers and other external consultants; 3) Adjudicative Providers: judges, arbitrators and their supporting organisations; 4) Non-adjudicative providers: mediators, conciliators and their supporting institutions; 5) Influencers: academics, government officers, policy makers': Herbert Smith Freehills and PWC (n 56) 6.

59 Ibid 20.

Finally, a recent survey[60] on cross-border disputes was conducted by the Singapore International Dispute Resolution Academy (SIDRA) of more than 300 international corporate executives, internal counsel and external counsel.[61] The survey found that international commercial arbitration remained the most popular dispute resolution choice amongst respondents, followed by international commercial litigation, hybrid processes (arbitration and mediation combinations) and then mediation.[62] These findings accord with perceptions of the dominance of arbitration and the relative light use of mediation in international commercial matters.

In terms of what factors drove respondents' preferences, the top-ranked responses were enforceability, neutrality/impartiality and cost.[63] The perceived lack of enforceability in respect of IMSAs may hence contribute to the lowly status of mediation's popularity relative to other processes. The SIDRA survey then examined specific responses of survey participants in relation to international commercial mediation. In this sense it is unclear as to how much experience respondents actually had with international commercial mediation. The most common reasons for selecting mediation were noted as 'impartiality of the forum', 'confidentiality' and 'speed'.[64] The authors note that '[un]like users of arbitration, mediation users do not rank enforceability very highly on their list of reasons to mediate',[65] attributing this view to a the lack of an internationally recognised, expedited enforcement mechanism for mediation. Thus, it is suggested that 'where compliance with the outcome of a dispute resolution process is a concern and enforcement mechanisms are a priority, mediation is less likely to be selected'.[66] The circularity in thinking evident here ignores the fact that participation in mediation may itself be enough to dispel the myth that post-mediation settlement enforcement is likely to be a problem. It is to this issue that we now turn.

2.3 Is the problem a real one?

Although the evidence is far from conclusive, the empirical research discussed above suggests there may be some demand from users, potential users, their advisors, and commentators, as well as current and putative mediation providers for the kind of uniform enforcement mechanism created by the Singapore Convention. Nonetheless, these studies at times suffer from a lack of precision as to who is making such claims. It is not always clear whether the views expressed are those of external lawyers, inside counsel, corporate clients, mediation providers or other commentators. Equally, it is not always apparent when such views expressed represent actual practical experiences in mediation or are merely observations. There may also be some distinct regional variations with some studies, for example, the SIDRA research, skewed towards Asia where the desire for formal regulation in the field may be culturally stronger.

---

60 Singapore International Dispute Resolution Academy, 'International Dispute Resolution Survey: Currents of Change' (Preliminary Report, 2019) <https://sidra.smu.edu.sg/sites/sidra.smu.edu.sg/files/documents/SIDRA2019_IDR_Survey_Preliminary_Report.pdf>.

61 In terms of breakdown, the study surveyed 194 'corporate users' (executives and in-house counsel) and 110 external counsel who had been involved in international commercial disputes between 2016 and 2018 from a range of jurisdictions, with around half of the respondents drawn from Asia: ibid 3.

62 Ibid 4.

63 Ibid.

64 Ibid 15.

65 Ibid.

66 Ibid.

In the international commercial context, there seems at best limited evidence that IMSAs are in fact not adhered to by the parties that crafted them.[67] The bulk of commentators accord with this view.[68] Although in the deliberations leading up to the Convention some delegates alluded to instances where settlement agreements had required cross-border enforcement proceedings,[69] as Schnabel notes, the main goal of the Convention is 'to provide an incentive to mediate in the many cases in which mediation might otherwise not be attempted. Ideally the Convention will rarely need to be invoked in court, as in most cases, parties will abide by the mediated settlements they conclude.'[70] Commercial settlements are generally undergirded with 'a "belt and braces" use of securities, guarantees and asset-backed obligations' to render them self-enforcing.[71]

Nonetheless, in terms of why endeavours to avoid compliance might take place, as one commentator put it: '[b]uyer's remorse, changing conditions over the course of the agreement's execution and other factors routinely can lead to parties to any negotiated settlement reneging their promise to execute'.[72] The study by Weiss and Griffiths did find some evidence of participants requiring at times to resort to litigation in respect of mediated settlements which had not been adhered to.[73] However, the basis for this proposition is weak.[74] It has also been noted by Nolan-Haley[75] in the US context that the highest number of litigated mediation cases (in recent years) concerned challenges to the enforceability of mediated agreements, and that generally challenges to mediated outcomes are on the rise. The US evidence may be skewed, however, by the increasing rise of court-based mandatory mediation in which recalcitrant parties, unwilling to attend mediation, may be more prone to trying to extricate themselves from mediated outcomes. Equally, Nolan-Haley suggests that part of the reason for this increase in challenges to mediated outcomes may be linked to the fact that 'in some contexts, the practice of mediation is becoming very much like the practice of traditional arbitration or like a judicial settlement conference', thus leading to greater prospect for parties to seek to unravel agreements.[76] This is important. Such trends, insofar as they currently exist, may in fact accelerate as a consequence of the

---

67 A 2015 study into international arbitration (which also touched upon mediation) found only 8% of respondents stating that they had experienced difficulties in enforcing IMSAs in the previous five years: White & Case, '2015 International Arbitration Survey: Improvement and Innovations in International Arbitration' (School of International Arbitration, Queen Mary University of London 2015) <http://www.arbitration.qmul.ac.uk/media/arbitration/docs/2015_International_Arbitration_Survey.pdf>.

68 As noted by Ahmed (n 7) 261: '[s]ettlement agreements resulting from mediation are more likely to be adhered to voluntarily'. Similarly, Mackie (n 22) suggests that 'few mediated settlement agreements fail to be implemented in the first place'.

69 See Intervention of Corporate Counsel International Arbitration Group (CCIAG), in UNCITRAL Audio Recordings: Working Group II (Arbitration and Conciliation, 62nd Session, 3 February 2015), cited in Schnabel (n 9) 4.

70 Schnabel (n 9) 4 (internal citations omitted).

71 James Maxey, 'Singapore Convention heralds new era for international mediation or is it a false dawn?' (*Buckles*, 8 March 2019) <www.buckles-law.co.uk/blog/singapore_convention_heralds_new_era_international_mediation>.

72 Birnberg (n 6) 2.

73 Weiss and Griffiths (n 51) 20–21 in which it was reported that of 92 respondents who answered this question 35% indicated that they infrequently had to relitigate mediated settlements that were not honoured with 8% indicating that they frequently had to do so.

74 The small numbers involved here and doubts around the extent that these respondents were speaking from significant experience mean that these findings should be treated with caution.

75 Jacqueline Nolan-Haley, 'Judicial review of mediated settlements: improving mediation with consent?' (2013) 5 Arbitration Law Review 152.

76 Ibid 153.

increased juridification of mediation likely to occur in the aftermath of the Singapore Convention. The authors return to this point below.[77]

It may be contended that uninitiated users or advisors participating in the surveys above are at times conflating their own experiences in arbitration with the promise of mediation. In arbitral proceedings it is the case, of course, that awards are imposed on parties. It may hence be natural that one party is unhappy and seeks to find ways of avoiding complying with its terms. Thus, against this backdrop, the New York Convention facilitating easier cross-border enforcement is an important measure in ensuring enforcement, and this backstop provides users with the confidence that the process will *de facto* be binding. In mediation, however, the parties are not bound to make agreements, and any such outcomes crafted are done so by the parties voluntarily. As has been widely noted, this is likely to lead to increasing 'ownership' of outcomes rendered, and the fact that mediated agreements are hence largely adhered to by parties without further enforcement activity has been proven to be the case in a range of contexts.[78] Other barriers to mediation's uptake are at play. These are discussed below.[79]

3 Might the Convention be counterproductive?

3.1 Profile raising

It could be argued that, even if the problem that the Singapore Convention seeks to tackle is largely illusory, it will do no harm and that the existence of the right to enforce outcomes will provide confidence to users.[80] Moreover, the resultant profile-raising of the international commercial mediation process shall significantly benefit its development. There is clearly some truth in this proposition. The profile of international commercial mediation has never been higher. There has been a glut of writing on the subject. Moreover, national and international conferences and webinars analysing the impact of the Singapore Convention are aplenty.[81]

The desire to adopt the Singapore Convention may be more pronounced in some jurisdictions than in others. As we noted above, the Global Pound Conference data suggested a greater preference for a uniform IMSA enforcement regime in some jurisdictions, such as Asia and Latin America, than in others, like Oceania and the UK. More generally, there may be more of a recognition culturally for increased regulation and formality as a proxy for legitimacy in some jurisdictions.[82] The Singapore Convention certainly responds to that audience.

---

77 See section 3.3 below.

78 See, for example, Craig McEwen and Richard Maiman, 'Mediation in small claims court: achieving compliance through consent' (1981) 18(1) Law and Society Review 11; Roselle Wissler, 'The effectiveness of court-connected dispute resolution in civil cases' (2004) 22 Conflict Resolution Quarterly 55, 65–68. Although, as the authors have already noted, increased lawyer domination of mediation may counter these effects.

79 See 4.1. below.

80 Veronika Vanisova, 'Current issues in international commercial mediation: short note on the nature of agreement resulting from mediation in light of the Singapore Convention' Research Paper No 2019/11/5 (Charles University in Prague Faculty of Law 2019).

81 See, for example, 'Singapore Convention seminar series' (*Immediation* 2020) <www.imimediation.org/2020/03/27/singapore-convention-online-seminar-series-the-line-up/>.

82 See, for example, Gloria Lim, 'International commercial mediation: the Singapore model' (2019) 31 Singapore Academy of Law Journal 377, 381, para 12, where it was noted that regulation in ADR, and mediation in particular, reflected aspects of Asian tradition and culture.

3.2 Operational difficulties

Some immediate problems spring to mind, however. At a basic level of acceptance, some major players such as the EU, the UK and Australia have not yet signed up to the Singapore Convention and could thus serve to stifle the aspirations of enthusiasts.

3.2.1 Wide scope of relief

Moreover, drafting issues may also have negative impacts. For example, as we noted above,[83] it can be argued that the scope afforded for potential challenges to settlements under the Convention is wide. In terms of the grounds for refusal to grant relief, their expansive nature has led to criticisms that this might lead to an undermining of the Convention. Some commentators, for example, have argued that excessive challenges may serve to counter important characteristics of mediation like confidentiality and evidential privilege.[84] Indeed, the authors argue that the mere existence of the Convention may *per se* lead to a greater *ex post* scrutiny of mediated outcomes and more attempts by parties to extricate themselves from outcomes rendered.

3.2.2 The Singapore Convention and creative settlements

One of the grounds under Article 5 by which a court or competent authority may refuse to grant relief is where an outcome is 'not clear or comprehensible'.[85] One significant potential benefit of mediation lies in the possibility of creative agreements.[86] By this we mean settlements beyond the gift of courts or common products of bilateral negotiations that can meet parties' underlying interests in a meaningful fashion. So, this kind of outcome could provide a framework for parties' future business conduct, including exploration of business interests and apologies tendered by one party to the other. However, not all such outcomes may be capable of being rendered in a definitive, legally robust fashion, and in this way may fall foul of the 'clear and comprehensible' exception.[87] As Phillips has stated: 'settlements (where mediated or not) result in agreements, with mutual obligations whose authority derives from the parties' consent, and often they are incapable on their face of being merely "enforced"'.[88]

Although it has been said that mediation's potential for creative settlements has been overstated,[89] a decline in this regard may be attributed to the rise of 'legal mediation' involving evaluative mediation practice and adversarial lawyering, leading to the narrowing of focus on legal matters and monetary settlements.[90] This trend may well continue in the post-Singapore Convention era with parties increasingly focusing on producing watertight, legally binding outcomes, especially when the existence of the Convention itself raises the issue of potential non-compliance more squarely into the parties' minds at the time of settlement formation.

83 See 1.4. above.

84 Phillips (n 21).

85 UNCITRAL (n 4) Article 5(1)(c)(i).

86 See, in the Scandinavian context, Lin Adrian and Solfrid Mykland, 'Creativity in court-connected mediation: myth or reality' (2014) 30(4) Negotiation Journal 421.

87 UNCITRAL (n 4) Article 5(1)(c)(i).

88 Phillips (n 21).

89 See, for example, Nancy Welsh, 'Making deals in court-connected mediation: what's justice got to do with it?' (2001) 79 Washington Law Quarterly 787, 813–816.

90 See Jacqueline Nolan-Haley, 'Mediation: The "New Arbitration"' (2012) 17 Harvard Negotiation Review 1.

3.3 The further juridification of mediation

Despite the generally flexible approach of the Convention, the authors contend that it may nonetheless change the underlying nature of international commercial mediation to render it the more natural terrain of the law and lawyers. There is a heightened role for lawyers anticipated by the Convention around, for example, providing for requisite formalities in and ensuring legally certain IMSAs, anticipating how agreements brokered may fare in different jurisdictions in which enforcement may be sought, and navigating the grounds for challenges to enforcement. As Mackie has noted, '[the Convention] is a win-win for mediators and lawyers. On the one hand, mediation now has further legitimacy in international trade rules. On the other, lawyers have an additional legal instrument on which they *can advise or seek grounds to argue on behalf of clients*.'[91] The idea that the Convention is a 'win' for lawyers is worrying. Any increasing role for lawyers in and around mediation should be viewed with caution. In this sense, much of the current debate around the increasing complexity and cost of international commercial arbitration is centred on the role of lawyers in the process.[92] Flood and Caiger tracked the growing historical development of the role of the lawyer and the law in commercial arbitration and the ensuing morphing of the process from a flexible, quick and business-friendly process to one which mimics formal legal process in many respects.[93] Similarly, Nolan-Haley laments the time when it was 'unnecessary to employ legal sanctions for compliance with arbitral awards because community members relied on each other's good faith' but that this no longer holds in an era in which 'arbitration was transformed into a more coercive and legalistic process'.[94]

The same charge has already been levelled at mediation.[95] Lawyers are, of course, important players in mediation. Their presence is often valued by mediators and their assistance in articulating claims, negotiating with an opponent and even protecting against mediator pressure is sought out by their clients.[96] The increasing domination of lawyers in the mediation process has, however, led to resultant rising costs, complexity and a narrowing of bargaining, so that there may be a greater focus on rights rather than interests. In a damning critique of the increasing legalisation of mediation in the US context, Nolan-Haley goes so far as to say that '[l]awyers generally control the mediation process, considering it the functional equivalent of a private judicial settlement conference'.[97] She also recounts the rise in evaluative mediators (often lawyers) behaving akin to arbitrators which has led to an increase in questionable unethical, adversarial lawyering and mediator 'spinning' by legal representatives in an attempt to influence such evaluations.[98] Moreover, research by Stipanowich in the context of district court

---

91 Mackie (n 22) (our emphasis).

92 See Steven Seidenberg, 'International arbitration loses its grip' (2010) 96 American Bar Association Journal 50, 55.

93 John Flood and Andrew Caiger, 'The juridification of arbitration: the juridification of construction disputes' (1993) 56 Modern Law Review 412 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3395051>. The dominance of lawyers as arbitrators is now a well-recognised phenomenon, see, for example, Thomas Stipanowich and Zachary Ulrich, 'Arbitration in evolution: current practices and perspectives of experienced commercial arbitrators' (2014) 25 Columbia American Review of International Arbitration 395, 404–405.

94 Nolan-Haley (n 90) 87–88.

95 Ibid.

96 Kathy Douglas and Becku Batagol, 'The role of lawyers in mediation: insights from mediators at Victoria's Civil and Administrative Tribunal' (2014) 40 Monash University Law Review 757, 784–785.

97 Nolan-Haley (n 90) 61.

98 Ibid 83.

mediation found that the offering of solutions or opinions by mediators had a long-term negative impact on participants' views as to satisfaction with these outcomes.[99] In a separate study he also found that mediators often felt 'bounced' into evaluation by lawyers.[100] In the Australian setting, Corbin et al have argued that there are still a number of lawyers who, within mediation, reject or at least find it hard to focus on client interests, sustaining relationships, and crafting solutions not centred on legal rights.[101] This view is supported by Rundell's research into court-connected mediation in the Supreme Court of Tasmania, finding that lawyers often held a preference for pursuing outcomes based on legal norms.[102] It has also been found that lawyers may adversely influence the ability of parties in mediation to reconcile with each other.[103]

The potential for affording greater prominence to the role of lawyers is especially troubling in the context of international commercial disputes. It has been argued that cross-border commercial mediation is already more legalistic in nature than its domestic counterpart, countering some of its inherent appeal. Mason, for example, has described international commercial mediation as one to which US companies bring more lawyers than they do to domestic mediation.[104] Similarly, Strong has noted that 'savings of time and money may be even less likely to occur in international commercial matters [than domestic mediation], where there is a tendency for counsel to conduct mediations like "mini arbitrations"'.[105]

Strong's empirical data suggest that parties are attracted to mediation in the cross-border context due to cost and time savings, as well as a desire for a more satisfactory process.[106] The recent Global Pound Conference research also points to users seeking out more collaborative, efficient approaches to dispute resolution.[107] It would be disappointing that if, in dealing with the perceived problem of enforceability, the *promise* of mediation's qualitative benefits is lost as more adversarial behaviour is supported through the increased prominence of lawyers in the process.

99 State Justice Institute and Maryland Judiciary, 'What works in District Court Day of Trial Mediation: Effectiveness of various Mediation Strategies on Short, and Long-term Outcomes' (Administrative Office of the Courts 2016).

100 Thomas Stipanowich, 'The international evolution of mediation: call for dialogue and deliberation' (2015) 46 Victoria University of Wellington Law Review 1191, 1211.

101 Lillian Corbin, Paula Baron and Judy Gutman, 'ADR zealots, adjudicative romantics and everything in between: lawyers in mediations' (2015) 38(2) University of New South Wales Law Journal 492, 503. See also Paula Baron, 'Throwing babies out with the bathwater – adversarialism ADR and the way forward' (2014) 40(2) Monash University Law Review 283.

102 Olivia Rundle, 'Lawyers perspectives on what is court-connected mediation for?' (2013) 20(1) International Journal of the Legal Profession 33, 52.

103 A study found that the presence of attorneys had a negative impact on the ability of parties to reach a reconciliation in their business relationship: Jean Poitras, Arnaud Stimec and Jean-Francois Roberge, 'The negative impact of attorneys on mediation outcomes: a myth or a reality?' (2010) 26(1) Negotiation Journal 9, 18–19.

104 Paul Mason, 'What's brewing in the international commercial mediation process' (2011) 66(1) Dispute Resolution Journal 64, 68.

105 S I Strong, 'Beyond international commercial arbitration? The promise of international commercial mediation' (2014) 45 Washington University Journal of Law and Policy 11, 16.

106 Strong (n 41) 2031.

107 Herbert Smith Freehills and PWC (n 56) 11.

4 Expediting mediation

4.1 Barriers to international commercial mediation uptake

In thinking about what measures might help expedite international commercial mediation, it is worth recalling the principal reasons why respondents to Strong's survey indicated that they would decline use. For those with experience of cross-border mediation, the main explanations included the lack of parties' and counsel's experience with mediation, fears about revealing litigation/arbitration strategy within the mediation, and concerns over mediator quality. Those without experience primarily cited a cultural preference for litigation or arbitration.[108] The attitudes of lawyers may be especially important. In the Global Pound Conference survey all stakeholder groups identified external legal advisors as the primary obstacle to change in commercial dispute resolution practice.[109] In the same survey, when lawyers (both external and in-house) were asked as to the most influential factors guiding their recommendations to parties about procedural options for commercial disputes, the most commonly cited reason was 'familiarity with the process'.[110] The lack of lawyer unfamiliarity then may be a major factor in stifling development of international commercial mediation.

These findings tally with more general ideas that the poor uptake of mediation across a range of contexts in a large part may emanate from the prevailing negative attitudes of lawyers and potential end users. Lawyers' resistance to mediation is well documented.[111] Such intransigence towards the process may emanate from ignorance and a lack of sophisticated appreciation of mediation, cultural dissonance or pecuniary interests.[112] This is important because (as suggested by the Global Pound Conference findings above) lawyers often act as gatekeepers to clients' involvement in mediation.[113] The winds are changing and clients, especially of the commercial variety, have of late become more powerful relative to their external lawyers and are typically more concerned with exacting efficiency of approach in their dispute resolution processes.[114] It may remain the case, however, that even sophisticated clients glean most of their knowledge of mediation and other dispute resolution processes from their lawyers.[115] The mediation option may rarely be raised as a possible course of action by disputing parties themselves.[116] It is also well known, however, that resistance to mediation may be a deep-seated phenomenon in the client base.[117] In this, mediation may be a hard sell. It speaks of compromise at a time of

108 Strong (n 41) 2034.

109 Herbert Smith Freehills and PWC (n 56) 3.

110 Ibid 3.

111 For a comprehensive review, see Bryan Clark, *Lawyers and Mediation* (Springer 2012) chapter 2.

112 Ibid.

113 For the classic treatment, see Douglas Rosenthal, *Lawyer and Client: Who's in Charge?* (Russell Sage Foundation 1974).

114 Julie Macfarlane, *The New Lawyer: How Clients are Transforming the Practice of Law* (2nd edn, University of British Columbia Press 2017) 138–139.

115 J Lande, 'Getting the faith: why business lawyers and executives believe in mediation' (2000) 5 Harvard Negotiation Law Review 137, 196.

116 See, for example, in the New Zealand context where it was noted that where client self-referral did arise, it was more likely to stem from the legal background of the client or where the client was a repeat player in mediation: Grant Morris and Amanda Lamb, 'Lawyers as gatekeepers to commercial mediation in New Zealand' (Resolution Institute/Victoria University 2016) 14–15.

117 See Craig McEwen, 'Managing corporate disputing: overcoming barriers to the effective use of mediation for reducing the cost and time of litigation' (1998) 14 Ohio State Journal on Dispute Resolution 1.

conflict. It asks parties to put their trust in a third party who is not their champion and will not adjudicate. The process cannot guarantee an outcome.

4.2 Dismantling the barriers

For domestic mediation, many endeavours in different jurisdictions to overcome weak growth of the process have been court-led.[118] Despite the fact that efforts to expedite development of cross-border mediation in the EU have also focused largely on court promotion endeavours,[119] it is not likely that such measures will be especially effective given the lack of involvement of the courts in international cases.[120] Other domestic measures mirror those that may help in the international commercial context. Numerous national jurisdiction reports have grappled with promoting mediation, including how sophisticated and meaningful awareness in the client and lawyer base can be raised to help grow the process, as well as help provide assurances over matters such as quality of mediation provision.[121]

4.2.1 Measures for lawyers and mediators

Absent court-led initiatives to promote mediation and force the hand of lawyers in the cross-border context, professional obligations may play a part. Indeed, these already exist. For example, in the EU, when lawyers are engaged in cross-border practice, they must act in accordance with the Code of Conduct for European Lawyers adopted by the Council of the Bars and Law Societies of the EU.[122] The Code provides that 'the lawyer should at all times strive to achieve the most cost-effective resolution to the client's dispute and should advise the client at appropriate stages as to the desirability of attempting a settlement and/or reference to alternative dispute resolution'.[123]

While education and training in mediation has grown enormously of late, there do exist some concerns regarding training standards and the diverse mediation practice norms found across jurisdictions. Linked to this issue is the fact that codes of practice for mediators vary significantly across different jurisdictions and that the use of sanctions applicable where codes are breached is patchy.[124] It has also been argued that expertise in domestic mediation is not necessarily a proxy for cross-border mediation ability given the

---

118 See, for example: Felix Steffek and Hannes Unberath (eds), *Regulating Dispute Resolution: ADR and Access to Justice at the Crossroads* (Hart 2013) 157.

119 Giuseppe De Palo et al, 'Rebooting the Mediation Directive: Assessing the Limited Impact of its Implementation and Proposing Measures to increase the Number of Mediations in the EU' (European Parliament, Directorate-General for Internal Policies 2014).

120 There are exceptions, including, for example, the promotion of mediation in the UK Commercial Court, where the substantial portion of cases heard are of the cross-border nature. See Judges of the Commercial Court in England and Wales (eds), *The Commercial Court Guide* (10th edn, HM Courts & Tribunals Service 2017) paragraph D8.9 <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/672422/The_Commercial_Court_Guide_new_10th_Edition_07.09.17.pdf>.

121 For a recent discussion of these issues in the UK context, see Bryan Clark, 'Where now for ADR?' (2019) New Law Journal <https://www.newlawjournal.co.uk/content/where-now-for-adr>; Bryan Clark, 'Some reflections on "I won't see you in court"' (2019) 2 Juridical Review 182.

122 Council of Bars and Law Societies of Europe, 'Charter of Core Principles of the European Legal Profession and Code of Conduct for European Lawyers' (2019).

123 Ibid 15, paragraph 3.7.1.

124 See the discussion in Katia Fach Gomez, 'The role of mediation in international commercial disputes: reflections on some technological, ethical and educational challenges' in Catharine Titi and Katia Fach Gomez (eds), *Mediation in International Commercial and Investment Disputes* (Oxford University Press 2019) <https://dx.doi.org/10.2139/ssrn.3418648>.

inherent complexity of the latter.[125] Reed, for example, has argued that 'the harsh reality is that, without the proper training, a first-rate domestic mediator is not naturally a successful international mediator'.[126] There may therefore be room for the further development of international standards through organisations such as the IMI that has already made significant strides in this field.[127]

4.2.2 In-house lawyers and contractual clauses to mediate

The Global Pound Conference data suggests that in-house counsel may be key facilitators in adopting mediation and in bridging the gap between clients and external lawyers.[128] Both their numbers and influence is rising. As Macfarlane has noted: 'the power and prestige of general counsel within any given corporation is greater now than ever before and reflects a desire to take control of outside legal fees and strategy'.[129] In comparison to external lawyers, in-house counsel may have none of the economic incentives that may tempt them to pursue certain forms of dispute resolution for their own interests. Nonetheless, they do share a cultural heritage and educational development with their external peers which may result in mediation being resisted.

One particular area where inside counsel can make a difference is through contract formation. At present the bulk of international commercial matters that find their way into mediation arise by way of a relevant dispute resolution clause in a contract.[130] As has been noted by Reed, '[t]he bottom line is that if in-house counsel are committed to saving time and resources with international commercial mediation, they must take the lead in drafting the necessary clauses into contracts, or instructing their external counsel to do so'.[131] Quek Anderson and Lee point to the discussions at the Global Pound Conference in this regard: 'several Panelists … spoke about inserting multi-tiered dispute resolution clauses into commercial contracts as a way of encouraging the business practice of attempting non-adjudicative processes first before proceeding to more drastic adjudicative forums'.[132]

4.2.3 Publicity of benefits

Direct selling to the client base is essential. In this, better promotion of success stories, in a language that potential users and in-house counsel understand, may help expedite use.[133] Equally, providing better guidance as to what the mediation shall entail may also be important, especially given the diversity of practice that can be encountered and the

125 Lucy Reed, 'Ultima Thule: prospects for international commercial mediation' Research Paper No 19/03 (NUS Centre for International Law 2019).

126 Ibid paragraph 3.1.

127 A detailed discussion of these issues is outside the scope of this article, but for an excellent discussion see Gomez (n 124) 14–19.

128 Herbert Smith Freehills and PWC (n 56).

129 Macfarlane (n 114).

130 See Mason (n 104) 68. The Singapore Convention itself, controversially, did not make provision for the uniform enforcement of agreements to mediate.

131 Reed (n 125) 9, paragraph 3.2. We do accept here that promoting mediation in this fashion may inevitably continue towards the increased juridficiation of mediation in that clauses may prescribe specific procedures for the mediation to follow. For example, jurisprudence in England and Wales points to the detailed procedural nature of such clauses required to ensure their enforceability: see *WAH (aka Alan Tang) v Grant Thornton International Ltd* [2012] EWCH 3198.

132 Dorcas Quek Anderson and Joel Lee, 'The Global Pound Conference in Singapore: a conversation on the future of dispute resolution' (2016) Asian Journal on Mediation 70.

133 Some useful case study examples can be found in Anthony Connerty, 'ADR as a "filter" mechanism: the use of ADR in the context of international disputes' (2013) 79(2) Arbitration 120, 128–133.

different cultural needs of potential participants. In the EU cross-border context, Lack and Schonewille have pointed to the diversity of practice in mediation arising from different national cultures and preferences, thus rendering it an inherently nebulous concept. One solution they advocate to help expedite cross-border commercial use is to educate parties as to the availability of the different species of the process so as to allow parties and their counsel to make an informed choice.[134]

Conclusion

It is understandable that the measures introduced by the Singapore Convention might be seen by enthusiasts as some kind of panacea to transform the relatively stagnant nature of international commercial mediation into a success story. Significant effort has gone into producing this new instrument and the hopes of many seem firmly pinned on it.

For the reasons put forward in this article, in the authors' view this trust is misplaced, however. In terms of its slow growth, the authors recognise that international commercial mediation is generally more complex than its domestic counterpart and, as a result, presents a range of additional hurdles to be overcome. Such matters include the multiparty nature of many cross-border commercial arrangements, cultural[135] and language barriers, and complex multi-contract arrangements.[136] Hence the speed and relative cheapness that advocates promote to potential users in domestic mediation may be more illusory in the cross-border context. Faced with this reality, mediation may be perceived as a tougher ask. Nonetheless, the promise of durability of outcome, preservation of relationships, flexible outcomes and process benefits remain. Further measures do need to be taken to unlock the potential fruits of mediation in the international commercial context. In the authors' view, however, tackling the perceived enforceability problem is not one of them. Indeed, there is a danger that the rush to embrace the Convention may lead to an undermining of international commercial mediation's essential attributes with the process all the poorer for it.

134 Manon Schonewille and Jeremy Lack, 'Mediation in the European Union and abroad: 60 states divided by a common word?' in Manon Schonewille and Fred Schonewille (eds), *The Variegated Landscape of Mediation: A Comparative Study of Mediation Regulation and Practices in Europe and the World* (Eleven International Publishing 2013) 20–21.

135 For a lucid examination of the potential problems of the traditional 'western' model of mediation in Asian jurisdictions such as Singapore, see Joel Lee, 'Culture and its importance in mediation' (2016) 16(2) Pepperdine Dispute Resolution Law Journal 317. See also Laurence Boulle, 'International enforceability of mediated settlement agreements: developing the conceptual framework' (2014) 7 Contemporary Asia Arbitration Journal 35, 52–53.

136 See Strong (n 105) 18–24; Reed (n 125) paragraph 2.1.

NOTES AND COMMENTARIES

NILQ 71(3): 503–522

The Supreme Court's judgment in *Adams* and the missing step of statutory construction

DAVID HORVATH-FRANCO*

University of Oxford

Abstract

The Supreme Court's recent decision in R v Adams *[2020] UKSC 19 was made partly on the basis of an assumption that the invalidity of the interim custody order made in respect of the appellant would automatically result in the quashing of his convictions for escaping detention on the basis of that order under paragraph 38(a) of Schedule 1 to the Northern Ireland (Emergency Provisions) Act 1973. However, to make this assumption is to skip a crucial step in the required reasoning: construction of the statutory offence the appellant was convicted of. Several arguments are put forward suggesting an alternative construction of paragraph 38(a). That construction holds that the paragraph 38(a) offence not only prohibits escape from detention under a valid interim custody order, but also from detention under an ostensibly valid, but nonetheless technically invalid, interim custody order.*

Key words: statutory construction; parliamentary intention; criminal law.

Introduction

On 13 May 2020, the Supreme Court gave judgment in *Adams*,[1] an appeal of two 1975 convictions for prison-escape attempts contrary to paragraph 38(a) of Schedule 1 of the Northern Ireland (Emergency Provisions) Act 1973 (the 1973 Act) in conjunction with the common law of attempts.[2] The defendant had been incarcerated under an Interim Custody Order (ICO) made ostensibly by the Secretary of State for Northern Ireland under Article 4(1) of the Detention of Terrorists (Northern Ireland) Order 1972. The challenge to these convictions was based on evidence gathered from government papers released under the 30-year rule,[3] namely advice given to the Attorney General in July 1974 to the effect that the ICO made in relation to the appellant may have been invalid. On the basis of that advice, the appellant argued that, since the Secretary of State had not personally issued the ICO, it was therefore void for unlawful delegation,[4] and, in turn, the offence of attempted escape was not made out on either count.

* With thanks to Alex Iordache for comments on an early draft. All remaining errors are my own.

1 *R v Adams* [2020] UKSC 19.

2 The criminal law of attempts in Northern Ireland was later codified by the Criminal Attempts and Conspiracy (Northern Ireland) Order 1983 (which in effect mirrored the Criminal Attempts Act 1981).

3 Section 5 of the Public Records Act 1958, as amended by section 1 of the Public Records Act 1967

4 *Carltona Ltd v Commissioners of Works* [1943] 2 All ER 560; *Vine v National Dock Labour Board* [1957] AC 488.

This line of argument includes an assumption that the invalidity of the appellant's ICO would necessarily result in his convictions under the 1973 Act being quashed. The Supreme Court, who found that the ICO was indeed void for unlawful delegation, also made this assumption; per the concluding paragraph of the judgment of the court:

> The making of the ICO in respect of the appellant was invalid. It follows that he was not detained lawfully. *It further follows that he was wrongfully convicted of the offences of attempting to escape from lawful custody and his convictions for those offences must be quashed.*[5] (emphasis added)

With respect to their Lordships, to make this assumption is to skip a crucial step in the reasoning required to deal with this appeal; namely, construction of the statutory offence of which (in conjunction with the common law of attempts) the appellant was convicted. This step is necessary in order to discover the essential elements of the statutory offence, and therefore also whether those elements were satisfied in light of the appellant's behaviour and other relevant circumstances.

Since there is no reference in the court's judgment to the 1973 Act at all, let alone to the words of the provision in question, it is doubtful whether the Supreme Court undertook such a process of construction. In fact, their Lordships' reference to the offence as that of 'attempting to escape from lawful custody' suggests an ignorance of the statutory basis of the convictions being appealed.[6] With respect to their Lordships, it is here submitted that their approach is therefore flawed, and so also that their substantive decision is open to criticism.

First, it will be shown why the step of construing the statutory offence under which the appellant was convicted is crucial in dealing with this appeal; second, it will be shown that it is at least arguable that, properly construed, the paragraph 38(a) offence does not require that the ICO under which the defendant is being detained be strictly valid. Consequently, the quashing of the appellant's convictions under paragraph 38(a) did not automatically follow from the finding of validity that their Lordships made with respect to the appellant's ICO.

---

5 [2020] UKSC 19, [41].

6 Ibid; see also [3], where the first count is again labelled as 'the offence of attempting to escape from lawful custody', but the second is referred to as 'a like offence', further indicating a lack of engagement on their Lordships' part with the precise basis of the convictions appealed. This oversight is not inherited from the Northern Ireland Court of Appeal's judgment below, where Sir Ronald Weatherup delivering the judgment of the court is very clear at [2018] NICA 8 [1]: 'This is an appeal against convictions on 20 March 1975 and 18 April 1975 on counts of attempting to escape from detention contrary to paragraph 38(a) of Schedule 1 of the Northern Ireland (Emergency Provisions) Act 1973 ('the 1973 Act') and common law. …'

1 Statutory construction as a necessary step

Where a case revolves around an Act of Parliament, two fundamental constitutional principles demand that it be properly construed by the adjudicating court.

The first is Parliamentary sovereignty.[7] Under this fundamental of British constitutional law, '[a]n Act can create any legal result'.[8] The consequence of this is that the words of a statute properly construed are capable of any meaning, notwithstanding any fundamental rights and other important principles of constitutional law (although the courts will naturally go to great lengths to uphold these).[9] In order to ascertain a statute's meaning, it is necessary to undertake a process of statutory construction, involving interpretation of the very words of the statute in light of their statutory, common law and other contexts.[10]

The second principle is the rule of law, under which public law powers (including that of prosecution and conviction) should be subject to strictly defined legal limits. This need is particularly heightened when a criminal statute is at play; the liberty (or at least reputation) of the defendant is at stake, and so the line between proscribed and non-proscribed action should be as clear as possible. This clarification is only achieved by direct reference to and explicit construction of the relevant statutory words. Such direct reference to the words of paragraph 38(a) is missing in the Supreme Court's judgment, which is entirely focused on the issue of the validity of the ICO imposed on the appellant at the time of his escape attempts. Had their Lordships engaged with the words of the statutory offence in question, it is submitted that they may well have found that it did not require that said ICO be valid.

---

7 *R (Miller) v Secretary of State for Exiting the EU* [2016] EWHC 2768 (Admin), at [20]: 'It is common ground that the most fundamental rule of UK constitutional law is that the Crown in Parliament is sovereign and that legislation enacted by the Crown with the consent of both Houses of Parliament is supreme.'; *R (Jackson) v Attorney General* [2005] UKHL 56, per Lord Bingham at [9]: 'The bedrock of the British constitution is ... the supremacy of the Crown in Parliament.'

8 Diggory Bailey and Luke Norbury, *Bennion on Statutory Interpretation* (7th edn, LexisNexis Butterworths 2017), section 2.2, page 25; see also A V Dicey, *An Introduction to the Law of the Constitution* (8th edn 1915) 38: '[Parliament has] the right to make or unmake any law whatever; and further, that no person or body is recognised by the law ... as having a right to override or set aside the legislation of Parliament.'

9 *Madzimbamuto v Lardner-Burke and George* [1969] 1 AC 645, Lord Reid at 723: 'It is often said that it would be unconstitutional for the United Kingdom Parliament to do certain things, meaning that the moral, political and other reasons against doing them are so strong that most people would regard it as highly improper if Parliament did these things. But that does not mean that it is beyond the power of Parliament to do such things. If Parliament chose to do any of them the Courts could not hold the Act of Parliament invalid.' However, cf Bailey and Norbury (n 8) section 2.2, page 25: 'The doctrine of parliamentary sovereignty does not, of course, mean that Parliament is free from conventional, political or practical constraints.' On balance, some interpretations are less likely (e.g. because they would entail a violation of fundamental rights), but they cannot be ruled out point blank. The court must inspect the words of the statute and only then can it rule out interpretations.

10 Philip Sales, 'Legislative intention, interpretation, and the principle of legality' (2019) 40 Statute Law Review 53, 56: 'Statutory meaning is conveyed through the medium of statements of law, rather than particular individual words. Individual statements, and individual words, have to be read in the context in which they appear. It is not sufficient to take meaning from a dictionary, though it may be relevant to do so. The context broadens out, from the immediate (the section itself), to the fasciculus of provisions or the Part of an Act in which the section appears, to the statute as a whole, to the wider legal context into which the section and the statute is inserted. The wider context also has varying degrees of remoteness from the exercise of legislative authority by parliamentarians. It includes the immediate legislative history, such as statements in White Papers, Law Commission reports or by ministers in Parliament, and then extends to more background matters such as the legislative forebears of the provision in question.'

2 Statutory construction of the offence

Statutory construction requires the court to go beyond the words of the statute to find in them the intention of Parliament. This is an objective intention, arrived at by reference to various principles of interpretation and against a background of relevant common and statute law.[11]

Although the Supreme Court neglected to do so, such an interpretive process should have been applied in *Adams* to the words of the statutory offence under which the appellant was convicted. A short demonstration of this process as applied to paragraph 38(a) to Schedule 1 of the 1973 Act is sufficient to show that the words of that provision are capable of more than one meaning, and so the offence proscribed therein of variable scope.

Paragraph 38(a) states that:

> Any person who … being detained under an interim custody order or a detention order, escapes … shall be liable on conviction on indictment to imprisonment for a term not exceeding five years or to a fine, or both.

A relevant principle to the interpretation of paragraph 38(a) is the principle of legality,[12] which holds that law must be clear, as well as ascertainable and non-retrospective. An interpretation of paragraph 38(a) in line with the principle of legality might be as follows:

> Any person who … being detained under **a valid** interim custody order or a detention order, escapes … shall be liable on conviction on indictment to imprisonment for a term not exceeding five years or to a fine, or both.

This is the interpretation implicitly adopted by the Supreme Court. Under this interpretation, the validity of the defendant's ICO or detention order is an essential condition of the offence, without which condition the offence cannot be committed.[13] This is, because if the ICO or detention order is not technically valid, it is not technically an ICO or detention order, and so there is no escape from detention under an ICO or detention order. As such, under this interpretation, since the appellant's ICO was invalid, both counts of attempted escape contrary to the paragraph 38(a) offence are not made out, and so both convictions are liable to be quashed.

However, other principles of interpretation may be relevant. Furthermore, paragraph 38(a) does not explicitly refer to a '*valid* interim custody order' or a '*valid* detention order',

11 *R v Secretary of State for the Environment, Transport and the Regions, ex parte Spath Holme Ltd* [2001] 2 AC 349, per Lord Nicholls of Birkenhead at 349: 'Statutory interpretation is an exercise which requires the court to identify the meaning borne by the words in question in the particular context. The task of the court is often said to be to ascertain the intention of Parliament expressed in the language under consideration. This is correct and may be helpful, so long as it is remembered that the 'intention of Parliament' is an objective concept, not subjective. The phrase is a shorthand reference to the intention which the court reasonably imputes to Parliament in respect of the language used. It is not the subjective intention of the minister or other persons who promoted the legislation. Nor is it the subjective intention of the draftsman, or of individual members or even of a majority of individual members of either House. These individuals will often have widely varying intentions. Their understanding of the legislation and the words used may be impressively complete or woefully inadequate. Thus, when courts say that such-and-such a meaning "cannot be what Parliament intended", they are saying only that the words under consideration cannot reasonably be taken as used by Parliament with that meaning.'

12 *Entick v Carrington* [1765] 95 ER 807, per Camden CJ: 'if this is law it would be found in our books, but no such law ever existed in this country'.

13 [2020] UKSC 19 at [41]: 'The making of the ICO in respect of the appellant was invalid. It follows that he was not detained lawfully. *It further follows that he was wrongfully convicted of the offences of attempting to escape from lawful custody and his convictions for those offences must be quashed*.' (emphasis added)

and so it is conceivable that the words of paragraph 38(a) might equally be read more widely as:

> Any person who … being detained under **an ostensibly valid** interim custody order or **an ostensibly valid** detention order, escapes … shall be liable on conviction on indictment to imprisonment for a term not exceeding five years or to a fine, or both.

Under this interpretation, the strict legal validity of the ICO/detention order is not an essential criterion of the offence; thus, despite the invalidity of the appellant's ICO, he might still be guilty of attempting to escape detention under it.

Of course, it may be that the Supreme Court saw it as obvious that the paragraph 38(a) offence would require any associated ICO or detention order to be valid at the point of commission of the offence. However, it is submitted that this is far from obvious; the two opposing interpretations of paragraph 38(a) shortly illustrated above are testament to this. Since more than one interpretation of paragraph 38(a) is clearly possible, their Lordships should have dedicated space in their judgment to the construction process in order to show that the interpretation that they implicitly adopted was the correct one. In order to further demonstrate the possibility of an alternative, wider construction, paragraph 38(a) will be discussed by reference to five canons of interpretation; namely, the mischief rule, the rule against doubtful penalisation, the principle of purposive construction, the principle of construction as a whole and the principle of effectiveness.

2.1 The mischief rule

According to Bailey and Norbury in *Bennion on Statutory Interpretation*:

> Parliament intends an enactment to remedy a particular mischief. It is presumed therefore that Parliament intends the court, in construing the enactment, to endeavour to apply the remedy provided by it in such a way so as to suppress that mischief.[14]

Authority for this approach to statutory interpretation, also called the 'mischief rule', can be found as far back as 1584, when the Barons of the Exchequer held:

> That for the sure and true interpretation of all statutes in general (be they penal or beneficial, restrictive or enlarging of the common law), four things are to be discerned and considered: (1) What was the common law before the making of the Act? (2) What was the mischief and defect for which the common law did not provide? (3) What remedy the Parliament hath resolved and appointed to cure the disease of the commonwealth. And (4) The true reason of the remedy; and then the office of all the judges is always to make such construction as shall suppress the mischief, and advance the remedy, and to suppress subtle inventions and evasions for continuance of the mischief, and *pro privato commodo*, and to add force and life to the cure and remedy, according to the true intent of the makers of the Act, *pro bono publico*.[15]

This approach has two limbs. First, the court must ascertain the mischief to which the statutory provision is addressed; mischief here doesn't just refer to wrongdoing committed by a defendant, but also to a normatively defective legal position which the

14 Bailey and Norbury (n 8) section 10.1, page 329'

15 *Heydon's Case* (1584) 3 Co Rep 7a.

statute aims to transform into a normatively superior one. Second, the court must then decide what Parliament intended to be the corresponding remedy; this is the means by which Parliament is to be taken objectively to have intended to transform the current legal position into the desired one.

In the case of paragraph 38(a), the second limb appears straightforward: the remedy which Parliament has introduced to resolve the mischief (whatever that may be) is to criminalise the conduct constituting the mischief. This is made crystal clear by the inclusion of the penalty: '[the defendant] shall be liable on conviction on indictment to imprisonment for a term not exceeding five years or to a fine, or both'.[16]

With regards to the first limb, it appears that the Supreme Court in *Adams* have accepted a narrow construction of the mischief to which paragraph 38(a) is addressed. This narrow construction is that the mischief addressed is the escape of people detained under either a *valid* interim custody order or a *valid* detention order. However, it is not clear that properly construed, the mischief at which the paragraph 38(a) offence is aimed is restricted to escape from detention under a strictly valid ICO or detention order. In fact, it is submitted that the order under which the defendant is detained at the time of his escape need merely be *ostensibly* valid. In other words, as long as a defendant, detained under an *ostensibly* valid detention order, escapes said detention, he commits the crime proscribed by paragraph 38(a).

This wider interpretation is advanced as it is submitted that the real mischief which Parliament is addressing in paragraph 38(a) is not the escape of a person lawfully detained, but rather interference with the UK's programme of internment in Northern Ireland.[17] Arguments to this effect are taken from the 1973 Act's statutory and historical contexts.

2.1.1 Statutory context

Useful context can be found in both the 1973 Act read as a whole and the background of statute law against which the Act was enacted.

The Act as a whole

The statutory context of paragraph 38(a) seems to support a wider construction. The 1973 Act gives the state wide-ranging powers in Northern Ireland, from a power to arrest without warrant on suspicion of terrorist activity[18] to an unlimited power of entry for 'preservation of the peace'[19]. It is clear that when the Act was passed, high levels of interference with fundamental rights, and punishment in the event of non-compliance, were countenanced. Of course, a court tasked with interpreting the Act would most likely do all it could to read fundamental rights into the statute, but such a reading-in is more difficult when one appreciates the Act's cumulative effect of trading fundamental liberties for safety and security. In order to maintain order and peace in Northern Ireland, Parliament in passing the 1973 Act granted wide powers of control over the population of Northern Ireland to agents of the state. The need to exert tight control over everyday

16 Paragraph 38(a) of Schedule 1 to the 1973 Act.

17 In an operation codenamed 'Operation Demetrious', between 9 August 1971 and 5 December 1975, 1874 Republicans and 107 Loyalists were detained without trial. See further Jonathan Bardon (1992) *A History of Ulster* (Blackstaff Press 1992); Paul Bew and Gordon Gillespie, *Northern Ireland: A Chronology of the Troubles 1968–1999* (Gill & Macmillan 1999); Tim Pat Coogan, *The Troubles: Ireland's Ordeal 1966–1996 and the Search for Peace* (Hutchinson 1995); David McKittrick, *Making Sense of the Troubles* (Blackstaff Press 2000).

18 Section 10(1).

19 Section 17(1)(a).

life in the province explains the high level of generality of these powers.[20] For example, section 6(1) gives police and the armed forces the power to stop anyone for the purpose of ascertaining their identity. If the person stopped refuses to answer, they commit an offence under section 16(2). Since the police and armed forces have no way of telling terrorist from ordinary citizen at first sight, they have a general power to question anyone; otherwise, their ability to control public areas would be seriously infringed. Likewise, the state's ability to control the movements of terror suspects through internment (detention without trial of suspected Irish Republican Army (IRA) members) would also be seriously infringed if the statutory support mechanism for internment – the paragraph 38(a) escape offence – were vulnerable to challenge based on the validity of the order authorising the defendant's detention. In summary, the mischief which paragraph 38 addresses is any interference with the state's ability to control the movement of terrorist suspects through internment. Furthermore, that mischief arises as much from escape from custody under an *invalid* ICO as it does from custody under a *valid* ICO.

This argument is not a novel one. It can be recognised in the reasoning of the England and Wales High Court in *The Queen v Davey*.[21] In that case, an order requiring the removal to hospital of a child suffering from scarlet fever had been made by a magistrate *ex parte* on the application of a local council official, and the child's mother, who refused to allow the child to be removed to hospital, was charged with obstructing the execution of the order under section 124 of the Public Health Act 1875. The defendant sought to collaterally challenge the order by way of defence to the prosecution. The Divisional Court held that she could not, Darling J finding that:

> ... the subject-matter with which this legislation [the Public Health Act 1875] deals is the spread of infection among people who have cases of infectious disease in their own homes; and if the proceedings to be taken by a local authority are of a dilatory character the legislation becomes nugatory. The object of the legislation is to get people so suffering into hospitals, and if the removal of the patients may be obstructed and the whole question of their removal argued *de novo* before the justices, the summary remedy is gone altogether, and the danger to the neighbourhood continues until the justices give their decision. Such a result was certainly never intended ... [22]

Darling J emphasised the importance of the local authority being able to act swiftly in order to combat the spread of infectious diseases. A parallel can be drawn between this case and *Adams*, the context of which is that it was vital that the authorities know the whereabouts and activities of terrorist suspects by bringing them into, and maintaining them under, custody. By inference, a further parallel can be made – that of the safeguarding against interference with a scheme of control. By obstructing the order made in respect of her child, the defendant in *Davey*[23] interfered with the council's scheme of control (of infectious diseases), and so was prosecuted for doing so – regardless of whether the order in question was valid or not. Likewise, regardless of whether the ICO made in relation to the appellant in *Adams*[24] was valid, by attempting to escape, he interfered with the state's system of control, as orchestrated through internment, and so committed the crime proscribed in paragraph 38(a).

20 See further below: '2.1.2 The historical context'.

21 [1899] 2 QB 301.

22 *The Queen v Davey* [1899] 2 QB 301, 304–305 (Darling J).

23 Ibid.

24 *R v Adams* [2020] UKSC 19.

Background statute law

Further support for this interpretation of the mischief aimed at by paragraph 38(a) is drawn from the background of statute law against which the 1973 Act was enacted. It is a principle of statutory interpretation that '[t]he text of an enactment must be read in its … legal … context'.[25] In the case of paragraph 38(a), it is arguable that the offence proscribed therein cannot be limited to escape from detention under a *valid* ICO, as that mischief is already adequately covered under section 26 of the Prison Act (Northern Ireland) 1953 (the 1953 Act), thereby making the paragraph 38(a) offence redundant.

Section 26(b) of the 1953 Act states that:

> Every person who … whether convicted or not, escapes from any prison or lock-up in which he is lawfully confined … shall be guilty of felony and shall on conviction thereof on indictment be liable to imprisonment for a term not exceeding three years.

Reading into the paragraph 38(a) offence a requirement that the relevant ICO be valid would equate to a requirement that the defendant's custody be lawful, thereby making the paragraph 38(a) offence come entirely under the scope of the section 26 offence.

Such an interpretation of paragraph 38(a) would therefore make the offence proscribed therein redundant, as the appellant could have been convicted on exactly the same factual basis under section 26 of the 1953 Act. Two principles operate against such a redundancy: first, the '[p]resumption that [an] enactment be given a purposive construction'[26] and, second, the 'general presumption against implied repeal'.[27] Provided they are not rebutted, these presumptions would lead to an interpretation of Parliament's intention as being to create a new offence in paragraph 38(a) of the 1973 Act, with a distinct scope from that in section 26 of the 1953 Act. Furthermore, it is submitted that this distinction hinges on the lack of a requirement that the defendant's custody had been strictly lawful.

On the other hand, it could be countered that Parliament in enacting paragraph 38(a) was not creating a new offence, but merely specifying the appropriate punishment ('a term not exceeding five years or to a fine, or both') for a particular factual context constituting escape under section 26(b) of the 1953 Act – namely, escape of someone lawfully detained under a valid ICO or detention order. However, two rebuttals may be made against this point. First, the contention that, in enacting paragraph 38(a), Parliament had wished to specify the appropriate punishment for escape in this factual context is weakened by the fact that Parliament did not specify the value of the fine to be dealt in the case of conviction. Second, elsewhere in the 1973 Act where Parliament has specified the appropriate penalty for another pre-existing offence, rather than creating a new offence altogether, this has been drafted significantly differently to paragraph 38(a). For example, section 1(1) of the 1973 Act, which replaces the death sentence for murder with life imprisonment in Northern Ireland:

> No person shall suffer death for murder and a person convicted of murder shall … be sentenced to imprisonment for life.

Significantly, Parliament in section 1(1) did not feel the need to codify the requirements for murder, as the subsection's sole purpose was merely to change the penalty for an existing crime. Therefore, if the sole purpose of paragraph 38(a) were clarification of the

25 Bailey and Norbury (n 8) section 9.2, page 287.

26 Ibid section 11.1, page 341.

27 Ibid (n 8) section 6.10, page 207.

appropriate penalty for a particular factual circumstance constituting escape under section 26(b) of the 1953 Act, one might expect it to be drafted in the manner of section 1(1), e.g:

> A person convicted of escape from detention under an interim custody order or detention order shall be liable to imprisonment for a term not exceeding five years or to a fine, or both.

Rather, the wording used in paragraph 38(a) is similar to that used in other offence-creating provisions of the Act, e.g. section 23:

> Any person who in a public place dresses or behaves in such a way as to arouse reasonable apprehension that he is a member of a proscribed organisation shall be liable on summary conviction to imprisonment for a term not exceeding six months or to a fine not exceeding £400, or both.

Parliament's intention in enacting paragraph 38(a) therefore seems to be offence-creating rather than punishment-specifying.

In consolidation of this argument, it is observed that, in addition to criminalising escape from detention under an ICO, the 1973 Act, under paragraph 38(b) of Schedule 1, also makes it a crime to

> … rescue … any person detained as aforesaid, or assist … a person so detained in escaping or attempting to escape.

It is important to note that not only was there already a crime of escape in Northern Irish law prior to the enactment of the 1973 Act, but also a crime of assisting escape. Section 29 of the 1953 Act provides that:

> A person who assists any person in escaping or attempting to escape from lawful custody, whether in prison or not, is guilty of an offence.

At face value the scopes of the two crimes under paragraph 38(b) of Schedule 1 to the 1973 Act and section 29 of the 1952 Act appear to be identical. Ultimately, both provisions criminalise the same conduct – assisting escape – and so as in the case of paragraph 38(a), it is arguable that in enacting paragraph 38(b), Parliament intended a difference in the elements of this new crime, i.e. that the detainee's detention need not have been strictly lawful.

Thus, the argument can be made that not only does Parliament in paragraph 38(a) create a new offence of escape whereby the lawfulness of detention is irrelevant, but in paragraph 28(b) also creates a new offence of assisting escape whereby the lawfulness of detention is likewise irrelevant. The inference can reasonably be made that Parliament's intention was to create a new set of offences mirroring those of escape and assisting escape already recognised in statute, yet without the requirement that the defendant's detention have been strictly lawful. This subtle distinction may not have been lost on those who prosecuted the appellant back in 1975, who elected to charge Mr Adams under paragraph 38(a) of Schedule 1 to the 1973 Act in conjunction with the common law, rather than under section 27 of the 1953 Act, under which:

> Every person who attempts to break prison or who forcibly breaks out of any cell or other place within any prison wherein he is lawfully detained or makes any breach therein with intent to escape shall be guilty of felony and shall on conviction thereof on indictment be liable to imprisonment for a term not exceeding five years.

	Escape	Attempted escape	Assisting escape
The detention must be lawful	Section 26(b) of the Prison Act (Northern Ireland) 1953	Section 27 of the Prison Act (Northern Ireland) 1953	Section 29 of the Prison Act (Northern Ireland) 1953
The detention need not be lawful	Paragraph 38(a) of Schedule 1 to the Northern Ireland Emergency Provisions Act 1973	Paragraph 38(a) of Schedule 1 to the Northern Ireland Emergency Provisions Act 1973 + The common law of attempts	Paragraph 38(b) of Schedule 1 to the Northern Ireland Emergency Provisions Act 1973

Table 1

Altogether, the set of mirroring offences may be grouped as in Table 1 above.

2.1.2 The historical context

Further evidence for construing the mischief targeted by paragraph 38(a) as interference with internment can be garnered from the historical context of the 1973 Act. That history is one of bloodshed and disorder.[28] On 30 March 1972, the Northern Irish government was dissolved and direct rule from Westminster imposed. Despite this measure, chaos and disorder continued in Ulster; the 'no go' areas of Belfast and Derry, physically retaken from the Provisional IRA by 12,000 British Army troops supported by tanks and bulldozers in July 1972,[29] remained firmly under Provisional IRA influence for the rest of the Troubles (and beyond). The lack of order in Northern Ireland had even brought the Troubles across the Irish Sea. Less than three weeks before the White Paper to the 1973 Northern Ireland Bill was first debated in the Commons on 28 and 29 March, the Provisional IRA had planted four car bombs in London, succeeding in exploding one outside the Old Bailey.[30] In light of this very recent historical context, there is a strong argument for interpreting Parliament's intention in passing the 1973 Act as first and foremost to exert overarching control over Northern Ireland and so, in turn, interpreting Parliament's intention in enacting the paragraph 38(a) offence as criminalising escape from custody under an ICO, regardless of the validity of that ICO (or at least before the ICO's validity could be determined under the proper procedure).

2.2 The rule against doubtful penalisation

As demonstrated above, application of the mischief rule might have led the Supreme Court to construe paragraph 38(a) widely, as covering defendants in custody under invalid as well as valid ICOs and detention orders. However, mitigating against the mischief rule, as well as such a wide interpretation of the paragraph 38(a) offence, is another important tenet of interpretation of criminal statutes with which their Lordships would have had to deal: the rule against doubtful penalisation.

28 See Jonathan (n 17); Paul Bew and Gordon Gillespie (n 17); Coogan (n 17); McKittrick (n 17).

29 Operation Motorman (31 July 1972).

30 The Old Bailey Bombing (8 March 1973).

The rule, recognised as being of constitutional importance by Lord Esher MR and Lindley LJ in *Tuck & Sons v Priester*,[31] and as stated in *Bennion on Statutory Interpretation*, is that:

> … a person should not be penalised except under clear law. The court, when considering, in relation to the facts of the instant case, which of the opposing constructions of the enactment would give effect to the legislative intention, should presume that the legislator intended to observe this principle. It should therefore strive to avoid adopting a construction which penalises a person where the legislator's intention to do so it doubtful, or penalises him or her in a way which was not made clear.[32]

Nonetheless, two counterarguments can be made against this point.

The first is that the rule does not apply in the immediate case, as the rationale underpinning the rule is not engaged here. That rationale seems to be that the scope of statutory offences must be clear enough to enable citizens to organise their lives around them. This principle is exemplified by *R v Dowds*.[33] In that case, the defendant alleged that section 2 of the Homicide Act 1957, as amended by section 52 of the Coroners and Justice Act 2009, now provided for acute voluntary intoxication as capable of giving rise to the partial defence of diminished responsibility on an indictment for murder. Referring to the above rationale, Hughes LJ dismissed the relevance of the rule to the case, saying that:

> There is simply no occasion which can be envisaged in which any citizen might order his affairs on the basis of a misunderstanding of the extent of the partial defence of diminished responsibility. The act of killing, with intent either to kill or to do grievous bodily harm and without justification (for example that of self-defence), must have taken place before there can be any question of the partial defence arising.[34]

Likewise, it is submitted that it is not realistic that a terror suspect, brought into custody under an ICO, would order his affairs on the basis of a misbelief that he will be immune to prosecution for escape if it can later be shown that the ICO is, for whatever precarious reason, in fact null. On the contrary, such behaviour might be described as 'chancing one's arm', a concept famously expounded by the American judge Oliver Wendel Holmes J:

> Wherever the law draws a line there will be cases very near each other on opposite sides. The precise course of the line may be uncertain, but no one can come near it without knowing that he does so, if he thinks, and if he does so it is familiar to the criminal law to make him take the risk.[35]

Glanville Williams said that 'people who chance their arms must take the consequences',[36] and this principle was endorsed by Thomas J (as he then was) in *Chigi v CS First Boston Ltd*.[37] Along these lines, it is argued that in attempting escape, the appellant in *Adams* chanced his arm. There is no indication that Adams knew that the ICO in relation to himself was vulnerable to challenge at the time of his attempted escape, so he chanced his arm in so doing. Even if he did suspect the ICO's invalidity, he still chanced his arm by attempting escape, since, short of a judicial review, he could not have been

31 (1887) 19 QBD 629, 638 (Lord Esher MR) and 644–645 (Lindley LJ).

32 Bailey and Norbury (n 8) section 27.1, pages 715–716.

33 [2012] EWCA Crim 281.

34 Ibid [38].

35 *US v Wurzbach* 280 US 396 (1930), 399.

36 Glanville Williams, 'Statute interpretation, prostitution and the rule of law' in C F H Tapper, *Crime, Proof and Punishment, Essays in memory of Sir Rupert Cross* (Butterworth 1981) 80.

37 [1997] All ER (D) 121.

sure that the ICO actually was invalid. Either way, by attempting to escape, Adams chanced his arm and so ought to take the consequences – i.e. a conviction under paragraph 38(a).

The second counterargument is that, even if the rule against doubtful penalisation is at play in *Adams*, it is not conclusive; per *Bennion*:

> In some cases … the court may find that the intention to impose the detriment was so strong as to require the doubt to be overridden.[38]

This approach is reflected in the case law. In *Bogdanic v Secretary of State for the Home Department*, Sales J (as he then was) warns that:

> [The principle against doubtful penalisation] is not an absolute principle. The overarching requirement is that a court should give effect to the intention of the legislator, as objectively determined having regard to all relevant indicators and aids to construction. The principle of strict interpretation of penal legislation is one among many indicators of the meaning to be given to a legislative provision. It is capable of being outweighed by other objective indications of legislative intention, albeit it is itself an indicator of great weight.[39]

The fact that the statute deals with an '[e]mergency' situation in Northern Ireland and the importance of maintaining physical control over terrorist suspects are examples of such indicators. Furthermore, Hansard contains several speeches by members of both Houses alluding to the need to re-establish control in Northern Ireland through the Act. The opening speech of the Bill's second reading in the House, delivered by the Secretary of State for Northern Ireland, Mr William Whitelaw MP, suffices to demonstrate this pressing need.

At the start of his speech, Whitelaw stresses that

> … there is no purpose to which [the Government] is 'more firmly committed than the restoration of the rule of law in Northern Ireland, and whatever means are necessary to that end will be made available'.[40]

Later in his speech, Mr Whitelaw refers to how

> … [i]t is the Government's intention that none of the provisions of the Bill, if it is passed, should continue in force a moment longer than it is needed.[41]

The Secretary of State also affirms that

> …[i]t remains the Government's intention that whenever possible people should be dealt with under the criminal procedures rather than those which may lead to detention.[42]

There seems to be at least the suggestion in the Secretary of State's comments that the government of the day was aware that the Bill's proposed detention scheme, as well as its other provisions, entailed curtailment of fundamental rights, but that this curtailment would be necessary in order to restore order to Northern Ireland, and that this and the other measures would only be used in so far as necessary to restore that order.

38 Bailey and Norbury (n 8) section 27.1, page 716.

39 [2014] EWHC 2872 at [48].

40 HC Deb 17 April 1973, vol. 855, col 276 <https://api.parliament.uk/historic-hansard/commons/1973/apr/17/northern-ireland-emergency-provisions>.

41 HC Deb 17 April 1973, vol 855, col 278 <https://api.parliament.uk/historic-hansard/commons/1973/apr/17/northern-ireland-emergency-provisions>.

42 HC Deb 17 April 1973, vol 855, col 285 <https://api.parliament.uk/historic-hansard/commons/1973/apr/17/northern-ireland-emergency-provisions>.

2.3 The principles of purposive construction, construction as a whole and effectiveness

A final argument for a wider reading of paragraph 38(a) to include escape from detention under an invalid ICO or detention order is one not focused on the elements of the offence itself, but rather on the possibility of challenging the validity of an ICO or detention order in criminal proceedings (as the appellant has done in *Adams*). This argument is made in light of the principles of purposive construction, construction as a whole and effectiveness. Specifically, the argument is that the presence of a statutory means of challenging ICOs and detention orders within the 1973 Act precludes collateral challenge of such an order in criminal proceedings. It follows therefore that the validity of the ICO or detention order under which the defendant was detained at the time of his escape should not be read into paragraph 38(a) as a necessary condition of the offence therein proscribed.

The 1973 Act's procedure for appealing ICOs is set out at paragraphs 26 and 27 to Schedule 1:

> **26**
>
> (1) Where a detention order has been made in the case of any person, he may within twenty-one days of the making of the order appeal by notice in writing to the Tribunal.
>
> (2) The Tribunal shall cause a copy of the notice of appeal to be sent to the Chief Constable and to the Secretary of State.
>
> **27**
>
> (1) A notice of appeal shall indicate the grounds of appeal and, where appropriate, the nature of any fresh evidence which the appellant wishes to tender on the hearing of the appeal.
>
> (2) Where notice of appeal has been given there shall be transmitted to the Tribunal a copy of the detention order and a copy of the record of the proceedings before the commissioner, which shall be in such a form as to indicate any part of the proceedings which took place in the absence of the appellant.
>
> (3) An appellant shall be entitled to receive a copy of the record of the proceedings before the commissioner, excluding any part of the proceedings which under paragraph 17 above took place in the absence of the appellant.

Per *Bennion*, the principle of purposive construction holds that:

> In construing an enactment the court should aim to give effect to the legislative purpose.[43]

The legislative purpose of paragraphs 26 and 27 is clearly to provide detainees with a means of challenging an ICO made in relation to them: the right to appeal the detention order is affirmed in paragraph 26(1), and the necessary steps to lodge an appeal laid out in paragraph 27(1).

Furthermore, according to the principle of construction as a whole:

> An act or other legislative instrument is to be read as a whole, so that an enactment within it is not treated as standing alone but is interpreted in its context as part of the instrument.[44]

43 Bailey and Norbury (n 8) section 11.1, page 341.

44 Ibid section 21.1, page 509.

Finally, per the principle of effectiveness:

> … the interpreter must construe the enactment in such a way as to implement, rather than defeat, the legislative purpose.[45]

The principles of construction as a whole and effectiveness require that an Act be interpreted holistically, in a way that gives effect to the Act's legislative purpose(s) and does not allow a legislative purpose found in one part of the statute to be undermined by another part. It is therefore necessary to interpret the rest of the 1973 Act (including paragraph 38(a) to Schedule 1) in a way that does not make the appeal procedure in paragraphs 26 and 27 redundant.

It is submitted that construing 'interim custody order' and 'detention order' in paragraph 38(a) as 'strictly valid interim custody order' and 'strictly valid detention order' would have that effect, as it would enable a means of challenging ICOs and detention orders extraneous to the statutory scheme. This alternative means would completely undermine the statutory scheme, as it would enable detainees to replace the risk of the Appeal Tribunal not deciding in their favour[46] with the risk of failing to escape, thereby obliterating any deterrent effect of the paragraph 38(a) offence. More than presenting attempted escape as a viable option, one might go as far as saying that interpreting paragraph 38(a) in this way would even encourage detainees to try to escape. If the attempt is successful, the detainee will have gained their freedom (however precarious that might be); if it fails, the detainee may still be able to gain their freedom in any event by collaterally challenging their conviction for attempted escape.

If this is so, as well as making the statutory appeal scheme in paragraphs 26 and 27 ineffective, such an interpretation would surely also be contrary to public policy, and so is not possible absent clear confirmatory words in paragraph 38(a).

2.4 Not too wide-ranging

It is submitted that the interpretation herein advanced of paragraph 38(a) of Schedule 1 to the 1973 Act is not too wide-ranging. However, it is understandable that it might at first glance appear to be so. A variation of the facts of *Adams* demonstrates this quite strikingly.

Imagine that back in 1973, Adams is taken into custody as per an ICO made in relation to him. Imagine then that when he asks to see this ICO, he is handed a piece of scrap paper, with the following words scribbled in crayon:

> ICO of Gerry Adams.
>
> Signed: FATHER CHRISTMAS

If we hold, as argued above, that paragraph 38(a) does not require that the ICO be valid, but merely ostensibly valid, and that the defendant has escaped from custody pursuant to it, then this Christmas-themed, yet clearly void, ICO is presumably sufficient to ground a prosecution if Adams attempts to escape. The ICO is ostensibly valid as it is being treated as valid by those holding Adams in custody. Surely this is too wide-ranging: the court might see the value in convicting Adams notwithstanding an invalidity of which he could not have known, but would likely be reluctant to attach criminal liability to a defendant who, handed an order apparently made by St Nicholas himself, is surely justified in thinking he is not in lawful custody.

---

45 Ibid section 9.8, page 310.

46 Paragraph 1 of Schedule 1 to the Northern Ireland (Emergency Provisions) Act 1973.

Two possible solutions to the problem are explored. The first is the adoption of a substantive/procedural invalidity distinction, under which some (substantive) invalidities will be fatal to the commission of the offence, whereas others (procedural) will not. The second is the inference into the paragraph 38(a) offence of a *mens rea* element as to the validity of the ICO or detention order. These two solutions will be discussed, and it will be shown that, while the first is ineffective, the second is quite satisfactory as a means of delimiting the paragraph 38(a) offence to within a reasonable scope.

2.4.1 Substantive/procedural validity distinction

The first solution – the adoption of a substantive/procedural invalidity distinction – was first endorsed by the High Court of England and Wales in *DPP v Bugg*.[47] In that case, two groups of defendants appealed against convictions for having entered military bases contrary to byelaw 2(b) of the RAF Alconbury Byelaws 1985, byelaw 2(b) of the HMS Forest Moor and Menwith Hill Station Byelaws 1986 and section 17(2) of the Military Lands Act 1892. Both sets of defendants raised the alleged invalidity of the byelaws for being made contrary to the requirements of section 17(1) of the 1892 Act as a defence to criminal proceedings in the Magistrates' Court. The Divisional Court held that the validity of a byelaw could be challenged in criminal proceedings if the invalidity resulted from a 'substantive' error, but where the error of law alleged was 'procedural', the byelaw remained effective until quashed in judicial review proceedings.

Woolf LJ described substantive invalidity as

> … where the byelaw is on its face invalid because either it is outwith the power pursuant to which it was made because, for example, it seeks to deal with matters outside the scope of the enabling legislation, or it is patently unreasonable. This can be described as substantive invalidity. … When the byelaw itself is alleged to be substantively invalid because of *Wednesbury* unreasonableness (*Associated Provincial Picture Houses Ltd. v. Wednesbury Corporation* [1948] 1 K.B. 223), for present purposes what has to be attacked is not the decision to make the byelaw but the byelaw itself.[48]

There seem to be two kinds of error described as 'substantive'; first, a straightforward error of law/illegality and, second, a *Wednesbury* irrationality error.

Woolf LJ in turn explained procedural invalidity as

> … where there is what can be described as procedural invalidity because there has been non-compliance with a procedural requirement with regard to the making of that byelaw. This can be due to the manner in which the byelaw was made; for example, if there was a failure to consult.[49]

Breaches of procedural fairness, therefore, such as the right to notice, would seem to come under 'procedural invalidity', as would a failure on the part of the decision-maker to comply with any statutory procedural requirements.

The substantive/procedural invalidity distinction could be adopted in the interpretation of paragraph 38(a) so that only substantive invalidity, but not procedural invalidity, of the ICO or detention order under which the defendant was being held could be raised as a defence. Accordingly, paragraph 38(a) to Schedule 1 of the 1973 Act would be construed thus:

47 *Bugg v DPP* [1993] QB 473.

48 Ibid 494 (Woolf LJ).

49 Ibid.

> Any person who … being detained under a **substantively valid** interim custody order or a detention order, escapes … shall be liable on conviction on indictment to imprisonment for a term not exceeding five years or to a fine, or both.

As such, substantive validity of the ICO is a necessary element of the offence, but procedural invalidity is not fatal to it.

The flaws in the procedural/substantive invalidity distinction

Notwithstanding its application in *Bugg*, there are significant problems with the procedural/substantive invalidity distinction, which was powerfully disapproved of by Lords Steyn and Irvine of Lairg in *Boddington*.[50]

One issue raised by their Lordships in that case was the distinction's incompatibility with post-*Anisminic* administrative law orthodoxy. Per Lord Irvine:

> In my judgment the reasoning of the Divisional Court in *Bugg*'s case, suggesting two classes of legal invalidity of subordinate legislation, is contrary both to the *Anisminic* case and the subsequent decisions of this House to which I have referred. The *Anisminic* decision established, contrary to previous thinking that there might be error of law within jurisdiction, that there was a single category of errors of law, all of which rendered a decision ultra vires. No distinction is to be drawn between a patent (or substantive) error of law or a latent (or procedural) error of law. An ultra vires act or subordinate legislation is unlawful simpliciter and, if the presumption in favour of its legality is overcome by a litigant before a court of competent jurisdiction, is of no legal effect whatsoever.[51]

The second issue is the unworkability of the procedural/substantive invalidity distinction due to its unclarity. Per Lord Steyn:

> There is also a formidable difficulty of categorisation created by *Bugg's case* [1993] Q.B.D. 473. A distinction between substantive and procedural invalidity will often be impossible or difficult to draw.[52]

Lord Irvine demonstrates this difficulty in his judgment:

> [In] my judgment the distinction between orders which are 'substantively' invalid and orders which are 'procedurally' invalid is not a practical distinction which is capable of being maintained in a principled way across the broad range of administrative action. This emerges from the discussion of *Wandsworth London Borough Council v. Winder* [1985] A.C. 461 by the Divisional Court in *Bugg v. Director of Public Prosecutions* [1993] Q.B. 473, 495-496. The court regarded it as a case of 'substantive invalidity,' i.e. in which either the decision to increase rents or the rent demands themselves were on their face invalid. I disagree. The rent demands appeared perfectly valid on their face. The decision was said by the tenant to be *Wednesbury* unreasonable, because irrelevant matters had, or relevant matters had not, been taken into account, as set out in his pleading. At trial, he would have had to adduce evidence to make out that case. It was not an error on the face of the decision. ... Many different types of challenge, which shade into each other, may be made to the legality of byelaws or administrative acts. The decision in *Anisminic* freed the law from a dependency on technical distinctions between different types of illegality. The law should not now be developed to create a new,

---

50 As outlined by Professor Forsyth: Christopher F Forsyth, 'Collateral challenge and the foundations of judicial review: orthodoxy vindicated and procedural exclusivity rejected' (1998) Public Law 364.

51 *Boddington v British Transport Police* [1999] 2 AC 143, 158 (Lord Irvine of Lairg LC).

52 Ibid 170 (Lord Steyn).

> and unstable, technical distinction between 'substantive' and 'procedural' invalidity.[53]

Thus, not only does Lord Steyn refute the workability of a substantive/procedural error distinction, but he also hints at a fundamental misunderstanding of the distinction between the two types of errors. Specifically, does the differentiating factor concern to what part of the decision-making process (procedure or substance) the error relates, or does it concern the patency or latency of the error? If the latter, then why use the terms 'substantive' and 'procedural'? The distinction itself seems to be imprecisely conceptualised, as well as difficult to make for certain errors.

A third issue with the distinction is an echo of the argument mentioned above regarding interpreting statutory offences in a way which would encourage contravention. The issue was raised by Lord Nicholls in *Wicks*, where His Lordship said:

> Further, it would seem to follow that in the case of procedural invalidity the defendant could be convicted even after the order was set aside as having been made unlawfully, so long as the non-compliance occurred before the order was set aside. In cases of substantive invalidity the citizen can take the risk and disobey the order. If he does so, and the order is later held to be invalid, he will be innocent of any offence. In cases of procedural invalidity, the citizen is not permitted to take this risk, however clear the irregularity may be.[54]

In some cases, therefore, a subsequent finding of invalidity will exonerate the defendant, whereas in others, it will make no difference. Thus, perhaps worse than simply encouraging citizens to risk committing crimes in contravention of challengeable administrative orders, the substantive/procedural invalidity distinction also arbitrarily discriminates between such risk-takers based on the type of alleged invalidity – a factor which is far-removed from the defendant's conduct, perhaps to the point of irrelevance.

In summary, since it is doctrinally dubious and both unworkable and arbitrary in practice, the substantive/procedural invalidity distinction should not be employed with respect to paragraph 38(a).

2.4.2 *Mens rea* requirement as to the validity of the ICO

The second, and preferred, solution involves introducing into the paragraph 38(a) offence a *mens rea*[55] element concerning the defendant's knowledge as to the ICO's validity. This would mean that the prosecution would have to prove that, at the point of his escape, the defendant either knew that the ICO/detention order under which he was being detained was valid, or that he was reckless/negligent as to its validity. This approach has two advantages. First, unlike the procedural/substantive distinction, a *mens rea* test is simpler, and any unclarity in its determination is left to the jury to decide as a question of fact. Second, a *mens rea* test fits better with the criminal law status of statutory offences as it focuses on the defendant's psychological guilt rather than the validity of an administrative act. Ultimately, paragraph 38(a) should be used to punish those who act criminally by negligently, recklessly or knowingly contravening ICOs/detention orders. Statutory

53 Ibid 159 (Lord Irvine of Lairg LC).

54 *R v Wicks* [1998] AC 92, 108 (Lord Nicholls).

55 David Ormerod and Karl Laird, *Smith and Hogan's Criminal Law* (14th edn, Oxford University Press 2015) 115: '*Mens rea* is simply the label for the mental element required by the definition of a particular crime – typically, intention to fulfil the elements of the actus reus of that crime, or recklessness whether they be fulfilled.'

crimes are presumed to have a mental element,[56] so determining that mental element is simply the next step for the court when interpreting the provision in question.

The expedience of this approach is evident when it is applied to *Adams*. Paragraph 38(a) would be presumed to include a fault element rather than be a strict liability offence, and so the interpreting court would then decide what exactly that fault element is. The Supreme Court was assumedly aware of one fault element: the appellant's intention to escape custody. However, the court does not appear to have considered a secondary fault element; namely, the appellant's state of mind vis-à-vis the validity of the ICO/detention order under which he was in custody at the time of his escape attempt. In fairness to the court, this fault element was not submitted to them by either party, although it perhaps went without saying that their Lordships, in construing a statutory offence, ought to have envisaged the possibility of a secondary fault element.

Assuming for the sake of argument that there is a secondary fault element to the paragraph 38(a) offence, and that that secondary fault element is recklessness as to the validity of the ICO or detention order, it could then be applied rather neatly to the facts of the case.

According to Ormerod and Laird, the recklessness 'test involves assessing two issues: (i) whether D foresaw a risk of the proscribed result or circumstances in the *actus reus* of the crime in question; and (ii) whether it was reasonable for D to take the risk in the circumstances known to him'.[57] With respect to Ormerod and Laird, limb (i) must be slightly adapted, as it is submitted that there need not be an exactly corresponding *actus reus* element (in this case, it is submitted that the strict validity of the ICO under which the appellant was detained is not an *actus reus* condition of the offence). Taking a risk that does not can be just as legally culpable as taking a risk that does; that is especially the case for paragraph 38(a), whose true mischief, it is submitted, is unjustifiable interference with the internment scheme. Such interference is done when a detainee escapes from detention, lawful or unlawful alike. Therefore, limb (i) of 'whether D foresaw a risk of a particular result or circumstances' is preferred.

There is no indication that Adams knew of the possible defect in the making of his ICO. That possible defect was recorded in a July 1974 opinion of JBE Hutton QC (later Lord Hutton of Bresagh), an opinion which only became public in July 2004. Furthermore, according to the Supreme Court's judgment, Adams only 'became aware of Mr Hutton's opinion in October 2009'.[58] Therefore, limb (i) is very clearly established; since the appellant had no reason to believe that his ICO was invalid, he very clearly foresaw a risk that it was valid. Furthermore, since he attempted to escape from prison thinking that he was being held under a valid ICO, he acted unreasonably in taking the risk that the ICO was valid, and so also satisfied limb (ii). Therefore, Adams satisfied the *mens rea* requirement of the paragraph 38(a) offence.

Now turning to the Christmas-themed scenario above, we can see how the additional *mens rea* requirement operates satisfactorily there to vitiate against an absurd result. If Adams had been taken into custody and, on asking to see the ICO authorising his custody, had been shown a piece of scrap paper signed in crayon by 'Father Christmas', a jury might conclude that, even if Adams still foresaw a risk that the crayon-written ICO was

56 Lord Reid in *Sweet v Parsley* [1970] AC 132, 148: 'whenever a section is silent as to mens rea there is a presumption that, in order to give effect to the will of Parliament, we must read in words appropriate to require mens rea'.

57 Ormerod and Laird (n 55) 129.

58 *R v Adams* [2020] UKSC 19 at [5].

valid (and thus satisfying limb (i)), he was still reasonable in taking the risk in assuming that it was a nullity, therefore not satisfying limb (ii). As a result, Adams would not have been reckless as to the validity of the ICO and the paragraph 38(a) offence would consequently not be made out.

This approach does in a way resemble the procedural/substantive invalidity distinction. For one, a substantive invalidity is more likely to be patent than a procedural one, and so a defendant who assumes that an act is substantively invalid is less likely to be found to be reckless in doing so than a defendant who assumes the presence of procedural invalidity. For example, the alleged *Carltona* defect in Adams – a procedural invalidity – was not at all clear to Adams until the release of government papers some 30 years later, whereas a decision so absurd as to satisfy the *Wednesbury* test – and so substantively invalid – might be sufficiently absurd for a defendant to assume it to be a nullity without being reckless as to its validity.

However, as demonstrated above, it is not always the case that procedural errors are latent and substantive errors patent, and so by cutting out the middle-man and directly focusing on the risk the defendant took in assuming the invalidity, the additional *mens rea* element approach does not fall into that conceptual quagmire, nor does it encourage citizens to risk offending.

Furthermore, whereas the procedural/substantive validity distinction suffers from issues both of doctrinal untenability and unclarity, there are no such issues with introducing a *mens* rea requirement as to validity of the challenged ICO. There is no clash with *Anisminic*, nor any need to engage in 'questionable'[59] categorisation of errors of law. In fact, any doubt as to the latency/patency of an error is integrated into the mental element of the offence and so, as a question of fact, left to the jury to decide. This approach also fits well with the decision of the House of Lords in *Head*.[60] In that case, the defendant had been convicted of having carnal knowledge of a mental 'defective' contrary to section 56(1)(a) of the Mental Deficiency Act 1913. However, it was accepted by both the England and Wales Court of Appeal and the House of Lords that the certificate of two doctors certifying that the victim was a defective and the Secretary of State's order transferring her to an institution were themselves defective and likely to be quashed. Consequently, the defendant's conviction was quashed. We might say that assuming the defendant's knowledge of the likely invalidity of the relevant orders concerning the mental state of the alleged victim, the defendant did not run a sufficiently high risk in ignoring the order so as to be reckless as to its validity, and so the quashing of his conviction was justified. The court might arrive at this result by finding that the strict validity of the order of the Secretary of State was not a necessary part of the offence under section 56(1)(a) of the Mental Deficiency Act 1913, but that recklessness as to its validity is a part of the *mens rea* of the offence, and that the defendant did not satisfy that recklessness requirement.

In conclusion, the reading into paragraph 38(a) of a *mens rea* requirement of recklessness as to the validity of the ICO, whose ostensible validity forms part of the *actus reus* of the offence, would be sufficient to abate any concern that the reading of paragraph 38(a) advanced here would be too far-reaching from a criminal-law perspective.

59 As described by Lord Steyn in *Boddington v British Transport Police* [1999] 2 AC 143 at 170.

60 *DPP v Head* [1959] AC 83.

2.5 The construction process – conclusion

In summary, considerable arguments can be proffered in favour of a wider interpretation of the paragraph 38(a) offence under which (in conjunction with the common law of attempts) the appellant was twice convicted. Specifically, this interpretation would hold that the strict validity of the ICO or detention order under which a defendant is detained at the time of his escape is irrelevant to his commission of the offence proscribed in paragraph 38(a). Consequently, it should not have been so readily assumed that the invalidity *simpliciter* of the appellant's ICO for unlawful delegation would result automatically in his convictions being quashed.

Conclusion

The aim of the article has been to highlight that the Supreme Court skipped a crucial step in their reasoning – the step of construing the very statutory provision under which the appellant was convicted. In fact, the article goes further than this, by using well-established canons of statutory interpretation to build a considerable case for a wider interpretation of the offence proscribed by paragraph 38(a) under which the validity of the defendant's ICO or detention order is not a necessary element of the offence. Regardless of whether the Supreme Court would agree with this interpretation (which it tacitly does not), by not explicitly engaging in the construction process, the court might be criticised for risking running afoul of both parliamentary sovereignty and the very rule of law which, ironically, many are bound to see their decision as upholding. Furthermore, by skipping the construction step, their Lordships forewent an opportunity to dispose of the case at a prior stage to determination of the lawfulness of the challenged ICO – a manoeuvre which might have allowed for a less politically charged conclusion to their judgment. Fault might more fairly be laid at the door of respondent counsel for not bringing arguments concentrating on proper construction of the 1973 Act to the forefront of their submissions; but it is certainly regrettable that a Supreme Court that has been in the past so willing and able to depart in its reasoning from parties' submissions when lacking would let itself be so restricted now.

NILQ 71(3): 523–530

Procurement and commissioning during COVID-19: reflections and (early) lessons

ALBERT SANCHEZ-GRAELLS*

University of Bristol

Abstract

This commentary reflects on some common themes that are starting to emerge in the early analysis of the healthcare procurement and commissioning response to the COVID-19 pandemic. Although it largely results from the observation of the situation in the English NHS, the most salient issues are common to procurement in other EU healthcare systems, as well as more broadly across areas of the public sector that have strongly relied on the extremely urgent procurement exception in the aftermath of the first wave of the pandemic. Given the disfunction and abuse of 'unregulated procurement' in the context of COVID-19, the commentary reflects on the longer term need for suitable procurement rules to face impending challenges, such as Brexit and, more importantly, climate change.

Keywords: procurement; commissioning; healthcare; COVID-19; pandemic; extreme urgency; unregulated procurement; probity; integrity; conflicts of interest.

Introduction

A few months after the start of the COVID-19 pandemic, the impact of the extreme public health emergency on public governance mechanisms is still being acutely felt, both internationally and domestically. The long-term implications of this crisis are yet to take shape,[1] but there are already some indications of potentially significant developments

* Professor of Economic Law and Member of the Centre for Health, Law, and Society at the University of Bristol Law School. This commentary results from active discussions in a webinar of the same title held at the University of Bristol Law School on 30 September 2020. I am indebted to the speakers and participants for very thought-provoking exchanges. A previous version of this commentary, including links to the presentations given by the speakers, was published in the University of Bristol Law School Research Blog on 5 October 2020 <https://legalresearch.blogs.bris.ac.uk/2020/10/healthcare-procurement-and-commissioning-during-covid-19-reflections-and-early-lessons-some-thoughts-after-a-very-interesting-webinar/>. Comments welcome: a.sanchez-graells@bristol.ac.uk.

1 For high-level reflection, see REFORM, 'Building a resilient state: a collection of essays' (October 2020) <https:/reform.uk/research/building-resilient-state-collection-essays>.

at EU level,[2] as well as evidence of the heightened risks for the UK's public health system post-Brexit.[3]

Healthcare procurement and commissioning have been at the very forefront of the initial reaction to the COVID-19 pandemic.[4] As discussed in an earlier contribution to this journal, where unforeseeable and extremely urgent circumstances not attributable to the contracting authority arise, public procurement rules get out of the way to free public buyers up to do all they can to get the required supplies and equipment.[5] The last few months have thus been, to put it simply, the longest period of (largely) unregulated procurement on record – and definitely the largest period since the inception of the EU public procurement rules.

The scattered evidence that continuously appears through the media,[6] as well as the more systematic reports that are starting to emerge,[7] show some evidence of positive adaptations to the immense challenges,[8] but also a clear disfunction of the (deactivated) mechanisms to ensure probity and integrity in procurement, and potential for abuse of extremely urgent 'unregulated procurement'.

Against this background, on 30 September 2020, the Centre for Health, Law, and Society of the University of Bristol Law School had the honour of hosting an excellent panel of speakers for a webinar on 'Healthcare procurement and commissioning during COVID-19: reflections and (early) lessons'. The speakers provided short presentations on a host of very complementary issues surrounding the reaction of NHS procurement and commissioning to the COVID-19 challenges. The ensuing discussion brought to light a number of general themes that are, by and large, aligned with the worries that others and

---

2 For discussion, see S Greer and A de Ruijter, 'EU health law and policy in and after the COVID-19 crisis' (2020) 30(4) European Journal of Public Health 623–624; T Clemens and H Brand, 'Will COVID-19 lead to a major change of the EU public health mandate? A renewed approach to EU's role is needed' (2020) 30(4) European Journal of Public Health 624–625; M Guy, 'Towards a European health union: what role for member states?' (2020) Journal of Risk Regulation, advanced access <https://doi.org/10.1017/err.2020.77>

3 See e.g. N Fahy, T Hervey, M Dayan et al, 'Assessing the potential impact on health of the UK's future relationship agreement with the EU: analysis of the negotiating positions' (2020) Health Economics, Policy and Law, advanced access <https://doi.org/10.1017/S1744133120000171>.

4 See e.g. Organisation for Economic Co-operation and Development, 'Stocktaking report on immediate public procurement and infrastructure responses to COVID-19' (updated 24 June 2020) <www.oecd.org/coronavirus/policy-responses/stocktaking-report-on-immediate-public-procurement-and-infrastructure-responses-to-covid-19-248d0646/>.

5 A Sanchez-Graells, 'Procurement in the time of COVID-19' (2020) 71(1) Northern Ireland Legal Quarterly 81–87.

6 For an early overview, see A Sanchez-Graells, '1 Billion problems in using extremely urgent public procurement to evade accountability?' (*howtocrackanut.com*, 17 May 2020) <www.howtocrackanut.com/blog/2020/5/18/1-billion-problems-in-using-extremely-urgent-public-procurement-to-evade-accountability>.

7 See e.g. House of Commons Public Accounts Committee, 'Whole of government response to COVID-19', Thirteenth Report of Session 2019–2021 (23 July 2020) paras 15–17 <https://publications.parliament.uk/pa/cm5801/cmselect/cmpubacc/404/40402.htm>; or National Audit Office, 'Investigation into how government increased the number of ventilators available to the NHS in response to COVID-19' (30 September 2020) <www.nao.org.uk/report/increasing-ventilator-capacity-in-response-to-covid-19/>. However, the more relevant reports are still work-in-progress. See e.g. National Audit Office, 'Supplying the NHS and adult social care sector with personal protective equipment (PPE)' <www.nao.org.uk/work-in-progress/supplying-the-nhs-and-adult-social-care-sector-with-personal-protective-equipment-ppe/>; and National Audit Office, 'Government procurement during the COVID-19 pandemic' <www.nao.org.uk/work-in-progress/government-procurement-during-the-covid-19-pandemic/>.

8 Confederation of British Industry, 'Public–private partnerships: lessons from COVID-19' (September 2020) <www.cbi.org.uk/media/5623/public-private-partnership-lessons-from-covid-19.pdf>.

I had expressed at the outset of the pandemic, and a number of challenges that will shape the readjustment or reregulation of procurement and commissioning in the medium and long term.

The remainder of this commentary initially provides some brief notes on the most salient points made by the speakers in their presentations, which do not aim to be exhaustive. It then goes on to offer my own reflections and views on what lessons can be extracted from the procurement and commissioning reaction to the first wave of COVID-19, which do not necessarily represent those of the panel of speakers.

1 Presentations by the panel[9]

Ms Neli Garbuzanova, Senior Procurement Manager (NHS) and Member of the European Commission Stakeholder Expert Group on Public Procurement, provided an 'Overview of the impact of COVID-19 on NHS procurement and commissioning'. Neli used a cycle-based analytical model of NHS procurement and commissioning,[10] highlighting the need for adjustments to the model in times of crisis but within certain limits. She presented the impact of COVID-19 at different stages of this cycle and stressed the various practical and governance challenges raised by the pandemic, as well as the main adaptations implemented during the first wave. Reflecting on potential governance shortcomings, Neli stressed the difficulties in enforcing rules on conflict of interest, as well as some of the new guidance on payment. She acknowledged the heterogeneous nature of pandemic 'needs', calling for a more targeted approach to commissioning and procurement with the lapse of time and increase of knowledge about the disease. Neli also offered interesting insights into future challenges derived from extreme urgency awards, including the need to modify 'rushed' contracts as the specific needs of a (potential) second wave become better understood. She also stressed the necessity for stronger collaboration at all levels, especially in times of crisis, which could build on some NHS national and local best practices.

Mr Rob Knott, Commercial Director, Digital (Guy's and St Thomas' NHS Foundation Trust) and former National Director of NHS Procurement, provided a very practical view of 'Procuring in times of COVID-19. Challenges at the frontline'. Rob provided an insiders' view on the tensions between centralised procurement and local pressures to ensure safety and continuity in frontline delivery of NHS services, as well as broader reflections on the challenges faced by the sudden and extreme change of circumstances that the pandemic brought. He made it clear that (further) digitalisation is very much needed in order to gain a better (almost) real-time view of operational needs and supply chain capacity to address them. He also emphasised how the pandemic stressed ongoing challenges in ensuring a proper functioning of centralised NHS procurement, and a more collaborative and integrated NHS provision. From a broader perspective, he also stressed that there will be an immediate need to pay more attention to issues such as embedding

9 The speakers' slides are available via the earlier blog post at the University of Bristol Law School Blog <https://legalresearch.blogs.bris.ac.uk/2020/10/healthcare-procurement-and-commissioning-during-covid-19-reflections-and-early-lessons-some-thoughts-after-a-very-interesting-webinar/>.

10 See N Garbuzanova, 'A dynamic whole-cycle approach to public procurement: a practitioner's perspective on best-practice methodologies' (2019) 6(2) African Public Procurement Law Journal 88–133. This is coupled with a companion website: <www.procurementart.com>.

social value[11] and emerging new approaches to procurement to support longer-term economic goals post-COVID.

Dr Pedro Telles, Associate Professor, Law Department, Copenhagen Business School, discussed the 'Scope and limits to the use of the negotiated procedure for extreme urgency grounds', building on the views he had previously published in his personal blog.[12] Pedro stressed that the reaction to the pandemic has evidenced the scope for potential abuse of the direct award of contracts on the basis of the 'extreme urgency' exemption, not only or primarily in healthcare, but across very significant areas of public sector activity. He reflected on the governance problems that arise if contracting authorities take a bullish approach to relying on this exemption from the procurement rules, as the ordinary checks and balances are unable to facilitate timely remedies in case of abuse. He went back to the basics of the EU and UK regulation of extremely urgent procurement and stressed the need for robust justifications of the proportionality and necessity of the 'unregulated' procurement, as well as for the lack of imputability of the circumstances generating the extremely urgent need to the contracting authority. He stressed in particular the need for cumulative analysis of all requirements.

Dr Kirsi-Maria Halonen, Adjunct Professor, Faculty of Law, University of Lapland, and Member of the European Commission Stakeholder Expert Group on Public Procurement, presented her reflections on the 'Governance challenges in extremely urgent procurement: do we need new rules?'. She provided further evidence of the scope for abuse of direct awards of contracts (in particular for personal protective equipment (PPE)) exempted on the basis of extremely urgent need. Kirsi pointed at some rather extreme cases emerging from the Nordic experience, which have even required the application of criminal statutes. This triggered her reflection on whether the current approach is adequate, as far as the existence of extreme urgency is largely understood to justify the complete setting aside of all integrity and accountability requirements. Her answer was a resounding 'no', which led her to advocate for a new minimum regulation of extremely urgent procurement that sought to preserve the application of crucial elements of the existing procurement framework, such as the application of mandatory exclusion grounds, a tighter control and neutralisation of potential conflicts of interest and the introduction of additional guarantees, for instance, where there are any advanced payments.

Dr Mary Guy, Lecturer in Law, Lancaster University Law School, and Founder of the Health in Europe – Virtual Discussion Forum, built on her insightful 2019 modelling of the interaction between the NHS and private providers of healthcare in England,[13] to reflect on the 'Changing patterns of patient choice under COVID-19. What longer term implications?'. Mary explained how the immediate reaction to the pandemic significantly disrupted the different channels for the provision of both NHS and private healthcare, in particular as a result of two collective agreements aimed to bring private capacity under the NHS umbrella to both support the management of the crisis and to facilitate continuity in the provision of key services. Mary also reflected on the underpinning exclusion orders, which disapply competition law in the healthcare sector for the purposes

11 See Procurement Policy Note on Taking Account of Social Value in the Award of Central Government Contracts (PPN 06/20, September 2020) <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/921437/PPN-06_20-Taking-Account-of-Social-Value-in-the-Award-of-Central-Government-Contracts.pdf>.

12 <www.telles.eu/>.

13 M Guy, 'Between "going private" and "NHS privatisation": patient choice, competition reforms and the relationship between the NHS and private healthcare in England' (2019) 39(3) Legal Studies 479–498.

of supporting that crisis and continuity-oriented collaboration between the NHS and private healthcare providers. She drew some interesting comparisons with similar approaches in other economic sectors (e.g. food retail distribution,[14] but also more generally[15]) and reflected on whether the situation in the healthcare sector is different and, as such, will perhaps lead to a longer-lasting differentiated treatment, raising questions about the long-standing relationship between the NHS and private healthcare.

2 Some reflections on common themes

There are a few common themes in the discussions surrounding the procurement and commissioning challenges that the COVID-19 pandemic brought. Some of these are hardly new issues, but the pervasiveness of the pandemic and the extreme economic burden of the procurement associated with it have perhaps brought a new sense of urgency in trying to find solutions (if there are any, or any practicable ones) to the integrity and efficiency challenges that extremely urgent procurement generates.

2.1 The limits of 'extreme urgency'

This is perhaps the obvious starting point. Extreme urgency is a sort of 'get away card' that sets aside procurement rules to free up public buyers to go to the market and get what is needed, when it is needed, (almost) no matter how.[16] There are important questions here, such as whether this is indeed the adequate approach, how long can this exemption apply for, or whether the exemption from 'standard' procurement rules should not imply the possibility of 'unregulated procurement', but rather the triggering of a lighter-touch regime that solely concentrated on key elements of probity and good commercial processes and practice. Ultimately, extremely urgent procurement and commissioning are still procurement and commissioning, in the sense that they have to be conducted through some process and by reference to some commercial practices. Therefore, there is no absence of regulation, but perhaps only informal regulation in that context.[17] The extent to which that should be formalised and constitute a special procurement regime will continue to trigger reflection and controversy. Equally, there are challenges in establishing the cut-off point at which the exemption can no longer apply because either the needs that are being satisfied through the relevant direct awards are no longer extremely urgent (e.g. in relation to foreseeable but not immediate needs), or because the continued extreme urgency is attributable to a lack of institutional reaction to the situation (e.g. in terms of dedicating additional resources, or prioritising some activities over others).

2.2 Is the healthcare sector (procurement) special?

This is another point of contention, as in particular it seems more difficult to insist on compliance with formal rules for healthcare-related procurement and commissioning when public health goals – i.e. lives – are directly at stake than, for example, in relation to the award of consultancy contracts of dubious direct relevance to the official messaging

14 On which see O Odudu, 'Feeding the nation in times of crisis: The relaxation of competition law in the United Kingdom' (2020) 19(2) Competition Law Journal 68–78.

15 O Odudu, 'UK & COVID-19: an overview of the competition policy and leading cases' (24 September 2020) e-Competitions, Article No 96054 <www.concurrences.com/en/bulletin/special-issues/uk-covid-19/uk-covid-19-an-overview-of-the-competition-policy-and-leading-cases-en>.

16 For more details, see Sanchez-Graells (n 6).

17 To avoid opening a can of worms on the concept of regulation, perhaps it is best to make a distinction between (binding) legal rules and other forms of regulation, such as guidance or practice. However, this issue exceeds the scope of this comment.

about social-distancing measures. Rather than the need for a particularly lenient approach in the healthcare context, which can result in the unjustified syphoning-off of large amounts of public resources to the wrong providers or the acquisition of useless kit, it is perhaps better to think about it in terms of justification for the exemption. There is a sense that the excuse of the pandemic has been abused in some corners to simply engage in cronyism, if not worse. So, the issue is not so much whether something concerns a need for the direct provision of healthcare services, but whether there is a sufficiently close link between the intended provision of services and relevant public health goals. It is also worth stressing, as clearly emerged from the discussions in the webinar, that proper procurement has public health value or, seen from the other end, that bad procurement generates negative public health outcomes (e.g. every time supplies need to be discarded due to not being fit for purpose, with the ensuing operational risks and economic waste).

2.3 If not procurement rules, then what?

Another common theme concerns the possibility of keeping a wide procurement exemption as is (or, rather, as wide as it can be) and rely on other regulatory tools instead – such as other public law controls on the expenditure of public funds and, where adequate, criminal law. To my mind, the alternative this seems to present is a false one. Under their 'normal operation', procurement rules do not deactivate the rest of public and criminal law rules, but rather seek to diminish the need to rely on them. What this usually creates is a certain atrophy of those mechanisms, as the proper running of the procurement system tends to minimise the cases needing those types of interventions – and as those mechanisms can be captured by political interests and manipulate the system for purposes other than ensuring the probity of public expenditure. This is exacerbated by the increasing reliance on the (judicialised) private enforcement of procurement rules, both as a mechanism to provide individual remedies and to uphold public interests, which seems to marginalise public oversight mechanisms to the uglier and uglier field of political football. From this perspective, it seems clear that taking all procurement rules out of the equation in the context of a major crisis can only put excessive pressure on a weakened set of accountability and responsibility mechanisms. This is not to say procurement rules should not be made more flexible in the face of extremely urgent needs, but rather to say that pinning all hopes of a proper and efficient expenditure of public funds on those other mechanisms may be wishful thinking. Once again, it seems that what is needed is more resource, including more resource for public oversight of compliance with the procurement and other rules in normal times, as well as in exceptional times.

2.4 The importance (and purpose) of the digital transformation

Another common theme in the discussions is that digital technologies can make a very important contribution to streamlining access to information and to facilitating an earlier intervention, in particular in supply-chain management and to adjust imbalances related to frontline operational needs. This concerns for example the need to have better and more readily accessible information about existing suppliers and about companies that seek to contract with the public sector for the first time in the context of an emergency – which will ring many a bell for those advocating for the creation of (centralised or interconnected) procurement registers – as well as a more granular and real-time view of stockpiles, goods in transit, and effective use of, for example, consumables. I find the thrust of this discussion very important because the main claim is not that the digital transformation should seek to replace current operational or decision-making mechanisms (although some tasks could clearly be automated), but rather seek to support

decision-makers and those at the frontline by giving them relevant, accurate and timely information on which to base their decisions. Similarly, the digital transformation should not seek to enforce rigid (algorithmic) determinations of supplier responsiveness, for instance, but rather empower contracting authorities to better assess their risks and, where appropriate, to take corrective or palliative measures (such as e.g. economic guarantees). All of this leads me to think that the digital transformation is in reality one new wave of information-based transformation and that, put in this light, it presents different opportunities and challenges than if thought of as the 'robofication' or 'AI-fication' of procurement and, for example, healthcare services provision.

2.5 An opportunity to get rid of systemic regulatory anomalies?

Another common theme is that the reaction to the pandemic could serve as a catalyst to correct systemic regulatory anomalies – e.g. the existence of convoluted 'standard' procurement rules that the 'unregulated COVID-19 procurement experience' could show as redundant or ineffective.[18] I am sceptical about the possibility of implementing significant changes, for instance, in the regulation of healthcare procurement, but also more generally. The reasons for this are difficult to articulate concisely, but suffice it to say that the pandemic is not precisely backing up the claim that unregulated procurement is superior to ordinary procurement in all cases and, as repeatedly stressed, the general approach to procurement regulation seeks to prevent bad procurement, rather than enable good procurement. While good procurement needs to be possible within the existing regulatory framework (and I think it is, as long as the flexibility of the system is properly understood and used), it seems clear that unregulated procurement simply generates the very risks that can undermine a proper functioning of the procurement function. Therefore, I think that it would be wrong to class procurement rules (or some/most of them) as an anomaly that gets in the way of commercial approaches by the public sector, in order to advocate for a much 'freer' procurement regime post-COVID. As others have argued,[19] and very much in line with what Kirsi-Maria Halonen advocated in the webinar, the lessons seem to go very much in the opposite direction and to take us back to basics.

3 Looking (not that far) into the future

Beyond the discussions in the context of the pandemic, some of the issues and central themes also relate to some challenges that lurk in the (not-too-distant) future. I am particularly worried about the following two.

3.1 Are there many new lessons to be learned? How hard should we try to identify them?

One of the issues that will require a difficult decision is the extent to which the current massive wave of extremely urgent procurement should be subjected to *post mortem* analysis and to what purpose – especially as effective remedies will be nigh impossible to obtain and political responsibility does not seem to be the currency of the times. This concerns both public/political scrutiny and academic scrutiny. As a point of principle, no stone

18 The same is said of e.g. Brexit, as there is a clear appetite to rid (UK) procurers of the existing controls on their exercise of commercial discretion; see e.g. S Arrowsmith, 'Reimagining public procurement law after brexit: seven core principles for reform and their practical implementation' Part I (http://ssrn.com/abstract=3523172) and Part II (http://ssrn.com/abstract=3672421).

19 L Folliot Lalliot and C R Yukins, 'COVID-19: lessons learned in public procurement. Time for a new normal?' (2020) 3 Concurrences 46–58 <http://ssrn.com/abstract=3685860>.

should be left unturned, on either front. However, this may not be a practical approach, or necessarily the preferable one, in particular concerning public oversight. Having already stretched public accountability institutions bogged down on the (myriad) details of COVID-related procurement and commissioning can detract from their ability to properly function post-COVID. This can be the case in the UK, for example, where the immediate future will bring the not smaller challenges of ensuring adequate governance, oversight and accountability of Brexit-related issues. In that regard, it can well be that a sort of 'amnesty' can be either formally declared or pragmatically adopted, so that regulators and other oversight bodies are not still looking at this crisis when the next one looms. Conversely, there seems to be a clear need for more (much more) academic research into the COVID-related practices and adaptations, both to draw the boundaries of any new 'light-touch regime' for (no longer unregulated) extremely urgent procurement and, perhaps more importantly, to reflect on how to maximise the opportunities brought forward by digitalisation from a governance perspective.

3.2 This is not the final crisis – what procurement rules for an ugly 'new normal'

The other thing that worries me, and much more, is that the pandemic is likely just a taster of the systemic distortions to come in our lifetime (not to mention the lifetime of our children and grandchildren). Given the institutional and social inertia against the adoption of truly radical sustainability-orientated procurement and consumption practices,[20] we can already foresee that dealing with the manifold implications of the climate emergency will at some point (rather soon) become the new (and from then on, permanent) extremely urgent need. At some point, extremely urgent will be the 'new normal' and that brings the question of what rules we need for that. From a dystopian perspective, it seems that unregulated procurement would only add corruption, maladministration and economic waste to the environmental meltdown – perhaps even accelerating the mutually reinforcing decomposition of institutional and physical ecosystems. The only way to try and strike a more positive chord (and try to avoid that ugly future) seems to be to get cracking with seriously rethinking procurement to bring sustainability (both environmental and economic/institutional) at its core, and to also get serious in harnessing the potential for digital technologies to accelerate its uptake.[21] In the end, there are no hopes for a vaccine against climate change.

Conclusion

This commentary has offered some reflections on the emerging themes and early lessons that can be learned from the procurement and commissioning response to the COVID-19 pandemic. These point clearly towards the need for more academic work on understanding the specific details of this response, as well as a broader reconsideration of the adequate regulation of extremely urgent procurement. It could well be that not many regulatory reforms are required and that the main improvements to be had depend on the proper harnessing of the potential that digital technologies offer. This area of exploration is particularly crucial in the face of the bigger challenges posed by climate change. It deserves our attention and effort.

20 For a call to action, see e.g. S Schooner and M Speidel, '"Warming Up" to sustainable procurement' (2020) 60(10) Contract Management 32–41 <http://ssrn.com/abstract=369742>.

21 For some exploratory thoughts, see A Sanchez-Graells, 'Digital technologies, public procurement and sustainability: some exploratory thoughts' (2019) SSRN working paper <http://ssrn.com/abstract=3482341>.

NILQ 71(3): 531–535

Pre-charge identification of a minor and Article 14 of the ECHR: *Judgment In the Matter of an Application by JKL (A Minor)*

FAITH GORDON

The Australian National University, Canberra, Australia

Introduction

> It is possible very effectually to poison the fountain of justice before it begins to flow. It is not possible to do so when the stream has ceased.[1]

Pre-charge identification of those accused of involvement in crime raises considerable concerns in relation to balancing the interests of all parties, while maintaining one of the core foundations underpinning the criminal justice system – the presumption of being 'innocent until proven guilty'.[2] The issue of pre-charge identification has been the subject of debate in recent years following the publication of the identity of a number of celebrities who were arrested in relation to allegations of historic sexual abuse and were subsequently released without charge.[3] Their arrests attracted intense media coverage, nationally and internationally.[4] A landmark decision following an action by Sir Cliff Richard against the BBC brought public and media attention directly to the consequences of pre-charge identification.[5]

There have been several calls for policy reform to this area of the law. In Northern Ireland, the negative consequences of pre-charge identification have been referred to by the Gillen Review, which explored the law and procedures in the context of serious sexual offences. The final report released in 2019 clearly recommended that: '[t]he identity of the accused should be anonymised pre-charge'.[6] However, previous attempts in the UK

---

1 *R v Parke* [1903] 2 KB 432, cited in C Burgess, 'Can "Dr Death" receive a fair trial?' (2012) 7(1) Queensland University of Technology Law Journal 16–28.

2 *Woolmington v DPP* [1935] AC 462, 481; Article 14 International Covenant on Civil and Political Rights. See also F Gordon, 'Guilty until proven innocent? How a legal loophole is being used to name and shame children' *The Conversation* (6 November 2017) <https://theconversation.com/guilty-until-proven-innocent-how-a-legal-loophole-is-being-used-to-name-and-shame-children-86073>.

3 Examples include: Paul Gambaccini, Matthew Kelly, Sir Cliff Richard and Jimmy Tarbuck.

4 For example: 'BBC radio presenter Paul Gambaccini arrested in sex abuse inquiry' *ABC News Australia* (2 November 2013); 'Cliff Richard's Berkshire property "searched by police in relation to alleged historical sex offence"' *Huffington Post* (14 August 2014).

5 *Sir Cliff Richard OBE v BBC* [2018] EWHC 1837 (Ch), Case No HC-2016-002849, 18 July 2018.

6 The Gillen Review <www.justice-ni.gov.uk/sites/default/files/publications/justice/gillen-report-may-2019.pdf>: Recommendation 10, page 30.

such as the unsuccessful Anonymity (Arrested Persons) Bill, introduced in 2011 by the Conservative MP and barrister, Anna Soubry, did not progress beyond second reading. On 21 January 2020 a version of the Anonymity (Arrested Persons) Bill [HL] 2019-21, went through a first reading in the House of Lords and, at the time of writing, a date for the second reading has not been scheduled.[7]

These existing discussions and unsuccessful policy interventions in the area of pre-charge identification have predominantly been centred on the experiences of adults who have been identified pre-charge and, notably, the same attention has not been given to the negative implications for minors who have also been identified pre-charge.[8] This altered in October 2015 with the arrest and police interview of a 15-year-old boy with Asperger's syndrome for an alleged cybercrime involving the 'hacking' of the databases of the TalkTalk telecoms company. The boy had his identity published, with details such as his name, age, place of residence and photograph, featuring in various media outlets including newspapers, such as the *Daily Mail*, *Daily Telegraph* and *The Sun*, as well as circulation via online media on Twitter and Google.[9]

1 Pre-charge identification of a minor

The case of *JKL* commenced in 2015, with proceedings in the High Court in Northern Ireland calling for the removal of the child's details and injunctions sought against future publications by organisations such as Google and Twitter.[10] In judicial review proceedings in 2016, counsel representing the child shone a light directly onto the lacuna in the current legislative framework.[11] Counsel for the applicant argued that the government's decision to implement legislation under Article 22 of the Criminal Justice (Children) (Northern Ireland) Order 1998, to cover reporting restrictions for children post-charge and at court, but not at the pre-charge stage for minors who are not charged with any criminal offence, was contrary to common law rules of fairness. It was also argued that failure by the Department of Justice in Northern Ireland to enact section 44 of the Youth Justice and Criminal Evidence Act 1999, introduced as part of New Labour's aim of modernising the youth justice system, was contrary to Article 8 of the European Convention on Human Rights (ECHR) and contrary to section 6 of the Human Rights Act 1998 (HRA).

Judge Colton, sitting in the Northern Ireland High Court, dismissed an application for judicial review in December 2016. In his judgment, Colton J acknowledged that, while the issue was one of social significance, he was of the view that the Department of Justice could not be compelled to legislate on this matter.[12] He asserted that the case made by the applicant relating to positive duties imposed on the state by Article 8 of the ECHR

---

7 For further details see Bill Stages – Anonymity (Arrested Persons) Bill [HL] 2019-21 <https://services.parliament.uk/Bills/2019-21/anonymityarrestedpersonsbill/stages.html>.

8 F Gordon, 'Pre-charge identification of minors: responses, rights and reform' (2020, forthcoming) Criminal Law Review.

9 'TalkTalk hack boy 15 arrested in Northern Ireland over attack' *The Independent* (London 26 October 2015) <http://www.independent.co.uk/news/uk/home-news/talktalk-hack-boy-15-arrested-in-northern-ireland-over-attack-a6709831.html>.

10 The applicant issued civil proceedings against Telegraph Media Group Ltd, Associated Newspapers Ltd, News Group Newspapers Ltd, Google Inc and Twitter International Company.

11 *Judgment in the Matter of an Application by JKL (A Minor) to Apply for Judicial Review and in the Matter of a Decision of the Department of Justice*: initial judgment delivered 21 December 2016 Ref COL10137 [2016] NIQB 99, and second judgment delivered 26 March 2020 Ref: COL11232 [2020] NIQB 29.

12 Initial judgment (n 11) paragraph 63.

was 'unsustainable',[13] and that a human rights challenge may not be brought on the grounds of a failure to legislate.[14]

2 Consideration of Article 14 of the ECHR

An appeal to the UK Court of Appeal saw the case remitted back to the High Court in Northern Ireland in 2019, for a first instance decision on a new argument raised by counsel for the applicant. The new argument centred on whether Article 14 of the ECHR had been breached.[15] In considering the question of discrimination, Colton J referred to *R (Stott) v Secretary of State for Justice*[16] and *R (DA & Others) v Secretary of State for Work and Pensions*,[17] and he directly applied Lady Black's four-stage approach in *Stott*. The four-stage test for Article 14 to be engaged comprises: the alleged discrimination must fall within the ambit of a Convention right; the different treatment must have been on the grounds of one of the characteristics listed in Article 14 or the applicant's 'other status'; the claimant and person who is treated differently must be in analogous situations; and there must be no objective justification for the different treatment.[18]

There was no dispute between the parties that the applicant's situation engaged his Convention rights under Article 6 and 8 of the ECHR.[19] The court considered whether the differential treatment complained of could be said to derive from the applicant's 'other status' under Article 14. Colton J considered the case law on the definition of 'other status', noting the difficulty in defining the concept as the 'jurisprudence as to what are the precise boundaries of "other status" is not clear'.[20] The decision in *Stott*[21] can be traced back to cases such as *Clift v UK*.[22] When the European Court of Human Rights considered the matter, it reviewed its decisions in which Article 14 was considered.[23] The analysis determined that the words 'other status' have been given a wide meaning[24] and that the court 'should take a "relatively broad view" of what constitutes "other status" in Article 14'.[25]

In the case of *JKL*, Colton J stated that '[w]hether or not the applicant enjoys "other status" for the purpose of Article 14 … is by no means straight forward' and that:

13 Ibid paragraph 28.

14 Section 6 of the HRA was drawn upon in this aspect of the case: ibid paragraphs 28–29.

15 Second judgment (n 11).

16 [2018] 3 WLR 1831.

17 [2019] UKSC 21.

18 *R (Stott) v Secretary of State for Justice* [2018] 3 WLR 1831, paragraph 8.

19 European Convention on Human Rights <https://www.echr.coe.int/Documents/Convention_ENG.pdf>. Article 6 of the ECHR outlines that: 'the press and public may be excluded from all or part of the trial in the interest of morals, public order or national security in a democratic society, where the interests of juveniles or the protection of the private life of the parties so require, or to the extent strictly necessary in the opinion of the court in special circumstances where publicity would prejudice the interests of justice'. Also, Article 8 provides for the right to respect for a person's 'private and family life … home … and correspondence'.

20 Second Judgment (n 11) paragraphs 25–27.

21 *Stott* (n 18).

22 [2010] ECHR 7205/07.

23 Second judgment (n 11) paragraph 31.

24 Ibid.

25 Ibid paragraph 41.

> the fact that he is a minor is innate or personal to him but his status for the purposes of this application depends on something extra, namely the fact that he has been arrested and interviewed in relation to a criminal offence.[26]

Colton J concluded that, in light of the 'liberal and broad interpretation' that has been adopted, the cause of the discrimination did fall within the remit of Article 14.[27] Colton J stated that counsel for the applicant had 'not really focused on whether the applicant is in an analogous situation to someone who appears before a court charged with a criminal offence but has rather focused on … what might be described as the pre-charge and post-charge dichotomy', namely whether there is 'lack of any objective justification for their different treatment'.[28] Counsel for the applicant proposed that any difference could not be justified. In response to this, counsel for the respondent put forward an argument that 'there is no reasonable expectation of privacy during proceedings in open court, as courts operate openly and subject to the full scrutiny of the public'.[29] However, '[t]he requirements of public justice are such that it is necessary to have statutory intervention to protect the interests of minors, particularly their Article 8 interests'.[30] Further, it was argued that there is no involvement of the court at the investigatory stage of the criminal process, and 'depending on the circumstances, a minor is likely to have a reasonable expectation of privacy at common law … through the law of privacy'.[31]

The court agreed that: '[g]iven the presumption of public justice the statutory intervention under Article 22 is necessary to protect the Article 6 and Article 8 rights of children who are brought before the courts', as '[o]therwise their identity will become public'.[32] It was asserted that the risk post-charge that a minor's identity would be made public derives from the principle of open justice, with media and public scrutiny of proceedings in court a present feature of open justice. The court observed that this 'is not the case with children who are in the "*pre-charge*" situation'.[33] Further, it was deemed that the risk of identification is significantly less, and, therefore, the applicant in this case was 'not in a relevantly analogous situation to children who actually appear before a court'.[34]

On 26 March 2020, the High Court held that the Department of Justice's failure to provide the applicant with pre-charge anonymity was not discriminatory and did not amount to a breach Article 14 of the ECHR.[35] In his judgment, Colton J concluded that he was satisfied that there existed a rationale behind the failure to commence section 44[36] and that any differential treatment had 'a legitimate aim'.[37]

26 Ibid paragraph 49.
27 Ibid paragraph 50.
28 Ibid paragraph 59.
29 Ibid paragraph 60.
30 Ibid.
31 Ibid paragraph 61.
32 Ibid paragraph 70.
33 Ibid. Emphasis in original.
34 Ibid paragraphs 70, 71.
35 Ibid paragraph 71.
36 Ibid paragraphs 71, 72.
37 Ibid.

3 The need for principled reform

Colton J has acknowledged that this case raises an important social issue. As argued elsewhere, the current media regulatory frameworks in the UK do not offer sufficient pre-charge protection for minors.[38] The international children's rights framework provides important requirements to ensure the protection of children from stigmatisation and further harm when they come into contact with the criminal justice system. For example, Article 16 of the UN Convention on the Rights of the Child 1989 states that among the guarantees for '[e]very child alleged as or accused of having infringed the penal law' is that 'his or her privacy [be] fully respected at all stages of the proceedings', and the UN Standard Minimum Rules for the Administration of Justice (the Beijing Rules) require a child's 'right to privacy' be 'respected at all stages' of the criminal justice process, 'in order to avoid harm being caused … by undue publicity or by the process of labelling', and 'no information that may lead to the identification of a juvenile offender shall be published'.[39]

However, in the *JKL* judgment the court's specific focus on risk of identification failed to consider identification and ramifications of identification for minors specifically. While the judgment appears to engage with rights under the ECHR, it does not appear to engage with these in a practical sense. Further to this, any discussion of the digital age is omitted. In the context of social media, there are central issues relating to permanency and third-party sharing which have not been explored. The case offered a key opportunity to address considerations of practical issues, but it failed to do so adequately. Rather, the court placed the onus on the common law to offer protection for minors through the law of privacy. In doing so, the court has failed to address the suitability of these available protections after the event of identification and after identity is known in the public domain. As this case has demonstrated, such existing protections are clearly not adequate.

Conclusion

This case is significant as it shines a much-needed light on the issues experienced by minors identified pre-charge and the lack of protections that exist. However, this case also represents a missed opportunity to address such significant issues. As the opening quotation demonstrates, it is possible to significantly negatively influence the 'foundation of justice' before it commences.[40] Thus, in this area of pre-charge identification, there is a direct and urgent need for legislative intervention and significant reform. The impact of the ever-changing environment of the digital age, with increased levels of third-party commentary and the permanency of the imagery shared on social media platforms, present significant challenges which need to be addressed by legislation, policy and regulatory frameworks in order to fully protect the rights of individuals pre-charge.

38 F Gordon 'Media regulation: strategies to mitigate the violence perpetrated against children who are publicly "named and shamed"' in W O'Brien and C Foussard (eds), *Violence against Children in the Criminal Justice System: Global Perspectives on Prevention* (Routledge 2020).

39 UN General Assembly 1985, rules 8.1, 8.2.

40 *R v Parke*, cited in Burgess (n 1).