Summer Vol. 74 No. 2 (2023)

NORTHERN IRELAND

LEGAL QUARTERLY

NORTHERN IRELAND
LEGAL QUARTERLY

JOURNAL INFORMATION

The *Northern Ireland Legal Quarterly* is a leading peer-reviewed journal that provides an international forum for articles, commentaries and notes in all areas of legal scholarship and across a range of methodologies including doctrinal, theoretical and socio-legal. The journal regularly publishes special issues within this broad remit.

Established in 1936, the journal has a history and rich vein of legal scholarship, combining distinct publications on the law of Northern Ireland, and prominence within the School of Law at Queen's University Belfast, with leading contributions to the discussion and shaping of law across the common law world and further afield. The School of Law at Queen's University Belfast took over the publication of the journal from SLS Legal Publications (NI) Ltd in 2008, where it has since been published quarterly. The journal became an online-only publication in January 2017.

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

AVAILABILITY AND ARCHIVES

The *Northern Ireland Legal Quarterly* is committed to making its contents widely available, to broaden our readership base. At least one article per issue is made available on an open access basis and may be published in advance. All articles become available on an open access basis on our website one year after publication.

All contributions to the journal become available on HeinOnline one year after publication (with issues going back to its launch in 1936) and LexisNexis three months after publication (with issues from 2019). The journal's contents appears on a growing range of indexing and abstracting services.

Since 2018 the journal's contents is promoted via social media and the Contributors' Blog.

In the summer of 2020, we expanded the reach and use of the *Northern Ireland Legal Quarterly* by adding 17 more years of content to the journal's existing archives. These now go back to 1999 (volume 50) and are widely accessed by our readership. Visit our Archive pages for further details.

SUBMISSIONS

The journal welcomes submissions of articles, commentaries, notes and book reviews on a rolling basis. Please see our 'For Authors' section for further details.

If you have any queries about the suitability of your article for the journal or if you have an idea for a special issue, please contact the Chief Editor Professor Mark Flear. For the contribution of commentaries and notes, please contact Dr David Capper. For book reviews, contact Dr Clayton Ó Néill.

SUBSCRIPTIONS

Subscriptions pay for a minimum of three months of exclusive access to the journal's latest contents (and up to one year for those who do not have access to LexisNexis).

NORTHERN IRELAND
LEGAL QUARTERLY

Summer Vol. 74 No. 2 (2023)

Contents

Articles

Commentaries and Notes

Northern Ireland
Legal Quarterly
Summer Vol. 74. 2 (2023) 211–236
Articles DOI: 10.53386/nilq.v74i2.1094

Death rites disrupted: coronavirus, 'lockdown' laws and the altered social ritual of the funeral

Heather Conway*
Queen's University Belfast

Correspondence email: h.conway@qub.ac.uk

ABSTRACT

While emergency measures for tackling coronavirus fundamentally altered our daily lives, this limiting of freedoms on public health grounds had an equally dramatic impact on the rituals of death. The sweeping restrictions imposed on the time-honoured social practice of the funeral recast its fundamentals but have not been meaningfully probed in legal scholarship. This article addresses that lacuna by examining the relevant laws and government guidance and their broader societal impact. Drawing on the multidisciplinary field of death studies, it examines both the transformative effect of these measures on funerals and the attendant human and social consequences. Integrating this analysis with evidence from emerging research on bereavement and grief during the pandemic, the article argues that the ongoing emotional toll of Covid-era funerals is fuelling a new type of public health crisis.

Keywords: coronavirus; lockdown laws; funerals; public health; rituals; grief.

Without tradition and ritual, death is a terrifying and meaningless experience.[1]

INTRODUCTION

Death, that one fate that awaits us all, is something that most people do not wish to contemplate. Modern Western societies are said to be 'death denying'; talking publicly about death and accepting its inevitability has become something of a social taboo. In his seminal work *The Hour of Our Death*, French historian Philippe Ariès attributed this

* Professor of Property Law and Death Studies, School of Law, Queen's University Belfast. I am extremely grateful to my colleagues, Professor Anne-Marie McAlinden, Dr John Stannard and Professor Anna Bryson, for their comments on an earlier draft of this article, and to the anonymous reviewer for their insights.

1 M H Jacobson and A Peterson, 'The return of death in times of uncertainty – a sketchy diagnosis of death in the contemporary "corona crisis"' (2020) 9 Social Sciences 131, 140.

twentieth-century phenomenon to specific factors that removed death from people's everyday lived experiences, from dying in hospital to the professionalisation of funerals.[2] Yet the global pandemic caused by the spread of SARS-CoV-2, the novel strain of coronavirus that emerged in late 2019,[3] made death a daily, and inescapable, fact of life. As cases surged across successive waves, and mortality rates climbed with unsettling and seemingly relentless speed, the worldwide devastation inflicted by a virus-clad Grim Reaper became all too apparent.[4]

The severe threat posed by the virus prompted radical public health controls in the spring of 2020, as individual governments devised strategies to tackle a rapidly evolving situation. The domestic legal architecture of the United Kingdom (UK) took the form of the Coronavirus Act 2020 and ancillary regulations[5] that fused a centralised approach with discrete powers for each of the devolved administrations.[6] Passed as time-limited, responsive measures designed to manage the effects of the pandemic and curb virus spread, the end result was an interventionist set of emergency rules supplemented by repeated government guidance and exhortations. Restrictions on movement, prohibitions on gatherings and social distancing were integral elements of a 'new normal', imposed on everyone as part of a collective effort to mitigate the Covid-19 crisis through radically different individual, group and societal behaviours. While these measures fundamentally altered our daily lives, the limiting of freedoms on public health grounds had an equally dramatic impact on the rituals of death. As part of lockdown rules and associated infection control measures, sweeping restrictions were imposed on funerals – restrictions that aggressively dismantled and recast the

2 P Ariès, *The Hour of Our Death* (Alfred A Knopf 1981). Ariès was one of a nucleus of scholars who formulated this narrative of denial, though the underlying thesis has been criticised. See M Robert and M Tradii, 'Do we deny death? I. A genealogy of death denial' (2019) 24 Mortality 247 and 'Do we deny death? II. Critiques of the death-denial thesis' (2019) 24 Mortality 377.

3 This article will use the generic term 'coronavirus' for this particular strain of the human virus. The term 'Covid-19' denotes the disease caused by the illness.

4 Even if this pandemic was not as deadly as others, such as the 1918 influenza pandemic that killed an estimated 25–50 million people globally.

5 The principal regulations were the Health Protection (Coronavirus, Restrictions) (England) Regulations 2020 (SI 2020/350); Health Protection (Coronavirus) (Restrictions) (Scotland) Regulations 2020 (SSI 2020/103); Health Protection (Coronavirus Restrictions) (Wales) Regulations 2020 (SI 2020/353 (W 80)); and Health Protection (Coronavirus, Restrictions) Regulations (Northern Ireland) 2020 (SR 2020/55). These regulations were subsequently repealed and replaced, with numerous ancillary regulations (too extensive to list individually) also passed at various stages of the pandemic.

6 This vast, hastily enacted legislative framework was not without criticism: see nn 15–16 below and accompanying text.

fundamentals of this time-honoured rite of passage, in a manner that few would have imagined pre-pandemic. It is these restrictions, and their societal impact, that form the basis of this article.[7]

Funerals, in the sense of a ritualised or ceremonial disposal of the body,[8] bring family, friends and communities together to mark the life of the deceased in an act of symbolic expression. As structured events where participants interact and engage in set patterns of behaviour, they create intensely personal yet simultaneously shared experiences for the bereaved that are, at once, both incredibly intimate and public.[9] Despite the breadth of law's involvement in death,[10] the prescriptive content of funerals has traditionally been shaped by a composite blend of personal choice, social convention, religious beliefs and cultural values; legal directives are kept to a minimum, beyond the mechanics of bodily disposal and environmental impacts.[11] However, all this changed with the emergence of Covid-19. As strictly enforced legal rules curbed attendance at funerals and promoted radically different behaviours, the funeral ritual as part of society's 'essential constitution', as something that 'speaks to people's core emotions and reveals values that a society holds dearest',[12] was transformed overnight. Until now, these measures and their influence on experiences of loss and mourning have scarcely been probed; while some legal scholarship has attended to the complexities arising from the effects of the pandemic

7 The social gathering of the wake, which is traditionally held before the funeral (though sometimes afterwards as a post-funeral gathering), is also an important element of the ritual of death. There are commonalities between these two sites of mourning in that close friends and family gather to remember the deceased, and to express their feelings. However, the focus here is on the funeral. Wakes are less formal, unstructured events that do not attract specific laws; and emergency coronavirus measures did not directly target wakes, beyond generic rules around numbers at gatherings and non-mixing of households.

8 Adopting the definition in T Walter, 'Bodies and ceremonies: is the UK funeral industry still fit for purpose?' (2017) 22 Mortality 194.

9 And regardless of variances between traditional, religious funerals and the more modern, secular ceremonies: see eg G Cook and T Walter, 'Rewritten rites: language and social relations in traditional and contemporary funerals' (2005) 16 Discourse and Society 365.

10 See H Conway, *The Law and the Dead* (Routledge 2016).

11 Isolated exceptions sometimes arise, where a particular funeral raises legality issues (eg a proposed open-air cremation in *Ghai v Newcastle City Council* [2009] EWHC 978 (Admin) and [2009] EWCA Civ 59) or public policy concerns (eg the funeral of Moors Murderer Ian Brady in *Oldham Metropolitan Borough Council v Makin* [2017] EWHC 2543 (Ch)).

12 G P Miller, 'The legal function of ritual' (2005) 80 Chicago-Kent Law Review 1181, 1181.

on funerary traditions,[13] the substantive issues identified in this article remain largely unexplored.

Although the subject has much wider resonance, this article looks at the legal response to Covid-19 in the UK, and its impact on embedded socio-cultural practices around death and dying, through the medium of the funeral.[14] The topic is significant for various reasons. While legal scholars busied themselves with a range of issues surrounding the emergency restrictions, from the deprivation of personal liberties[15] to the impact on social care systems and mental capacity laws,[16] the impact on the funeral as a basic human institution was curiously neglected. The pandemic not only gave death a captive, global audience; it was also the first time that the funeral as a universally recognised and carefully constructed ritual[17] was hastily reconfigured by law and public health directives – something that assumed even greater importance when higher mortality rates throughout the pandemic made funerals a lived reality for more people. In stimulating a much-needed debate on the significance of this particular paradigm shift, the article makes another distinct contribution by fusing legal analysis with insights gained from the multidisciplinary field of death studies. Seeing the funeral as a blend of organic rituals that performs specific tasks for both the living and the dead[18] not only contextualises the legal issues. It also reveals

13 See eg H Conway, 'Funerals and coronavirus in Northern Ireland: new legal rules, new social norms' (2020) 2 Folio: Northern Ireland Conveyancing and Land Law Journal 16 and C Nyamutata, 'Funerary rites and rights of the dead: jurisprudence on Covid-19 deaths in Kenya, India and Sri Lanka' (2023) 12 Global Journal of Comparative Law 36.

14 The funeral itself (as the dispositive element) is one in a series of multiple events that mark an individual's death: eg visiting the deceased's family, displaying the body, the wake or post-funeral gathering. Commemorative events, such as memorial services and the scattering of ashes post-cremation, are other key elements. Emergency restrictions affected all these things, though the current focus is on the funeral.

15 See eg C Nyamutata, 'Do civil liberties really matter during pandemics? Approaches to coronavirus disease (Covid-19)' (2020) 9 International Human Rights Review 62; J Pugh, 'The United Kingdom's Coronavirus Act, deprivations of liberty, and the right to liberty and security of the person' (2020) Journal of Law and the Biosciences 1; and T Hickman QC, E Dixon and R Jones, 'Coronavirus and civil liberties in the UK' (2020) 25 Judicial Review 151.

16 I Antova, 'Disability rights and Covid-19: emergency laws and guidelines in England' (2020) 28 Medical Law Review 804 and A M Farrell and P Hann, 'Mental health and capacity laws in Northern Ireland and the Covid-19 pandemic: examining powers, procedures and protections under emergency legislation' (2020) 71 International Journal of Law and Psychiatry 101602.

17 See G M Bosley and A S Cook, 'Therapeutic aspects of funeral ritual: a thematic analysis' (1994) 4 Journal of Family Psychotherapy 69.

18 The core literature is referenced in the third part of this article: 'What funerals "do" and the impact of coronavirus restrictions'.

the transformative effect of these public health measures, and the attendant human consequences.

The first part of the article examines the traditional public health narrative around the fate of the dead (bodily decay) and explains how the recently deceased posed a different sort of contaminant risk to the living through community transmission of coronavirus at funerals. Having detailed the main restrictions, the article notes the public health fixation with physical health and not contracting Covid. This is exposed as being selectively and temporally myopic because it ignored the emotional impact of the emergency restrictions, something that is very evident in the funeral context. Drawing primarily on the fields of sociology, psychology and anthropology, the next section examines the role that funerals play in saying goodbye to the dead and caring for the living by fostering connectedness, providing essential social support and facilitating the grieving process. The article goes on to assess how this was changed by the process of law-making during the pandemic; it also probes interpretative issues surrounding the emergency measures, the difficult choices faced by the bereaved, and posits reasons for widespread compliance with the rules. Finally, the article considers the impact of disrupted death rites on individuals and communities. Drawing on a growing body of work that connects lockdown funerals to emotional trauma and complicated grief, it sets out a rudimentary framework of analysis for a new and emerging public health crisis.

PUBLIC HEALTH IMPERATIVES AND DEALING WITH THE DEAD

The law's treatment of human remains has always been based on two core values: respect for the dead, and public health fears around decaying corpses. The first is a universal standard that permeates the law.[19] Respecting the dead not only speaks to basic notions of human dignity; as a society, and as individuals, we care about how our dead are treated because it gives us a sense of existential comfort about our own treatment when we die. The second speaks to the threat of disease as unattended bodies decompose: to guard against this, and the risk of sensory or visual offence, the dead must be physically separated from the living. These core values counter-balance each other in a delicately poised set of virtual scales, as twin imperatives that death laws must

19 From judicial statements to this effect (see eg Martin J in *Calma v Sesar* (1992) 106 FLR 446, 452: the dead must be 'disposed of without reasonable delay, but with all proper respect and decency') to international humanitarian laws mandating the respectful treatment of those killed in armed conflict (see eg article 16 of the Fourth Geneva Convention).

reflect and respect. However, in pandemics and other emergencies,[20] the balance inevitably tips heavily to one side.

Addressing a 'serious and imminent threat to public health'[21] was the overriding objective of the legislative framework and government messaging for tackling coronavirus.[22] While the legal response was unparalleled, the containment measures relied on orthodox suppression techniques of prevention and containment;[23] constant edicts to wash our hands, practice social distancing and self-isolate were supplemented by legal rules prohibiting social gatherings and placing restrictions on movement to curb person-to-person transmission.[24] In using emergency measures to protect public health, the working assumption was that all citizens posed a significant risk of virus spread. And while the primary focus was on the living, the dead featured strongly in the public health narrative, though not in the conventional sense.

A different contaminant risk?

The contaminating potential of the corpse was the formative basis of a raft of burial laws passed in nineteenth-century England when a toxic combination of urban expansion, unsanitary conditions, and high mortality rates overcrowding graveyards posed a major threat to public health.[25] Fast forward to Covid-19, and the contaminant risk of the dead being in proximity to the living was still part of the socio-legal narrative, though not through the natural process of decay. At the start of the pandemic, concerns around the infectious properties of corpses and possible virus transmission through respiratory droplets and bodily fluids necessitated protocols for handling human remains in suspected

20 Eg natural disasters.

21 This was the justificatory language used in the' Introductory text' to the first set of principal regulations passed in March 2020 (see n 5 above), and in subsequent iterations of the core regulations across the four nations: see eg the Health Protection (Coronavirus, Restrictions) (No 3) (England) 2020 (SI 2020/750) and the Health Protection (Coronavirus) (Requirements) (Scotland) 2021 (SSI 2021/277).

22 As part of this, the Coronavirus Act 2020 augmented existing public health measures in the Public Health (Control of Disease) Act 1984, and extended specific powers under the 1984 Act (which applies to England & Wales) across the UK.

23 See A Wilder-Smith and D O Freedman, 'Isolation, quarantine, social distancing and community containment: pivotal role for old-style public health measures in the novel coronavirus (2019-nCoV) outbreak' (2020) 27 Journal of Travel Medicine.

24 The specific legal rules are discussed below.

25 M E Hotz, 'Down among the dead: Edwin Chadwick's burial reform discourse in mid nineteenth century England' (2001) 29 Victorian Literature and Culture 21.

or confirmed Covid-19 deaths.[26] This necessitated pathologists and funeral directors wearing personal protective equipment, bodies being placed in closed coffins, and both long-standing social traditions and religious or cultural rites being halted where these involved contact with the body (including washing or viewing the corpse).[27] However, a greater danger quickly emerged: that of the dead inadvertently acting as highly localised sites for person-to-person transmission, by bringing the living into close contact with each other through the medium of the funeral and, for example, shaking hands or hugging to express condolences.[28] This public health risk is what drove the restrictions imposed on *all* funerals in late March 2020, for both Covid-19 and non-virus deaths.

Emergency measures and the reconfigured funeral

The Coronavirus Act 2020 was silent on the issue of funerals, focusing instead on systems management, and ensuring that national and regional authorities had capacity to handle dead bodies with care and dignity as mortality rates increased.[29] Funerals were dealt with under the statutory regulations, introduced across the four nations of the UK within days of the 2020 Act becoming law.[30] Operationalising specific powers conferred by the Act,[31] the initial result was a fairly uniform set of restrictions in England, Scotland, Wales and Northern Ireland.

---

26 Though the level of risk was unknown: see L G M Dijkhuizen, H T Gelderman and W Duijst, 'The safe handling of a corpse (suspected) with Covid-19' (2020) Journal of Forensic and Legal Medicine 1.

27 See eg E Crubézy and N Telmon, 'Pandemic-related excess mortality (Covid-19), public health measures and funerary rituals' (2020) 22 eClinicalMedicine 100358. These restrictions had a huge impact on various religious and ethnic communities, given the centrality of the prohibited acts to their death rituals: see eg 'The Muslim bereaved cruelly deprived of closure by coronavirus' *The Guardian* (London 7 July 2020).

28 As occurred in Africa during Ebola outbreaks: C Park, 'Traditional funeral and burial rituals and Ebola outbreaks in West Africa: a narrative review of causes and strategy interventions' (2020) 5 Journal of Health and Social Sciences 73.

29 Pandemic preparedness, and expediting burial and cremation processes (if necessary), necessitated a number of measures. These included changes to death certifications (2002 Act, ss 18–21 and sch 13) and inquests (ss 30–32), and ensuring that potentially large numbers of dead bodies could be transported, stored and disposed of with respect (s 58 and sch 28).

30 The discussion immediately below focuses largely on the 'principal regulations', passed in late March 2020 and listed at n 5 above, which triggered the initial changes to funerals. For brevity, for subsequent amendments to these regulations, and additional or replacement measures, selective examples across the four nations are used.

31 In particular, the power to issue directions in relation to events, gatherings and premises under s 52 and sch 22 of the 2020 Act.

Respect for the dead and bereaved families meant that funerals could take place with mourners present,[32] while the basic dispositive options of burial or cremation remained available.[33] Funeral directors were permitted to operate as essential services,[34] while places of worship, crematoria and burial grounds that were (initially) closed to the public could open for funerals,[35] which were a listed exception to the prohibition on public gatherings of (what was originally) two or more persons.[36] And while people were prevented from leaving their homes without 'reasonable excuse', one such justification was to attend the funeral of '(i) a member of the person's household, (ii) a close family member, or (iii) if no-one within … (i) or (ii) [was] attending, a friend'.[37] Government guidance on funerals and public health more generally supplied further granularity, imposing numbers caps by limiting attendance to a maximum number (initially 10 people both indoors and outdoors – though slowly increased across the four nations under subsequent statutory regulations)[38] and insisting on

32 Unlike Italy, where high mortality rates and strict confinement of citizens to their homes meant that funeral services (both religious and civil) were banned early in the pandemic: 'Coronavirus: how Covid-19 is denying dignity to the dead in Italy' (*BBC News Online* 25 March 2020).

33 An earlier version of the Coronavirus Bill that appeared to allow forced cremation (assuming a surge in deaths and lack of grave space) was hastily amended, following a backlash from Members of Parliament (MPs) and faith groups: see 'Emergency coronavirus legislation altered after Muslim and Jewish communities raise concern over forced cremation' *The Telegraph* (London 23 March 2020).

34 As businesses exempt from closure and other specific restrictions: reg 5(1) and sch 2, pt 3 of the principal regulations in England; reg 6(1) and sch 1, pt 4 in Wales; reg 4(1) and sch 2, pt 3 in Northern Ireland; and reg 4(1) and sch 1, pt 3 in Scotland.

35 See regs 5(5), (6) and (8) of the principal regulations in England; regs 7(1) and (3) in Wales; regs 4(5), (6) and (8) in Northern Ireland; and regs 4(6), (7) and (9) in Scotland. However, social-distancing measures and operational constraints meant that some buildings could not, or chose not to, allow any mourners: see the examples listed at n 74 below.

36 See reg 7(c) in England; reg 8(5)(c) in Wales; reg 6(c) in Northern Ireland; and reg 6(c) in Scotland.

37 See regs 6(1) and (2)(g) in England; regs 8(1) and (2)(g) in Wales (though the Welsh Government also added a 'carer' of the person attending as a fourth category); regs 5(1) and (2)(g) in Northern Ireland; and regs 5(1) and 8(4)(g) in Scotland.

38 In line with different restrictions on gatherings, which saw eg England move to a maximum of 30 mourners in April 2020 under the Health Protection (Coronavirus, Restrictions) (No 2) (England) Regulations 2020 (SI 2020/684). It appears that Northern Ireland was the only jurisdiction to set numerical limits *for funerals* in statutory regulations, in the first months of the pandemic: see Conway (n 13 above) 17–20.

strict social-distancing measures throughout.[39] Persons from different households would have to travel separately to and from the funeral, and sit or stand at least two metres apart from each other, something that later became a legal mandate,[40] alongside the requirements to wear face coverings in enclosed spaces (introduced in the summer of 2020)[41] as the living were spatially segregated to minimise the risk of virus spread.

The minutiae of these funeral restrictions[42] changed periodically, as the pandemic ebbed and flowed.[43] However, the two core elements – the macro-level changes that altered the social fabric of the funeral more than anything else – remained mercilessly intact for over a year. Upper limits on mourners and physical spacing created a new, state-imposed governance structure; these, in effect, became terms and conditions that the bereaved had to accept when burying or cremating their dead. A core theme of what follows, though, is an assertion that the same restrictions on funerals were an overreach on public health grounds.

Public health and selective myopia

Going back to the critical phases of the pandemic and looking at public health through the analytical prism of coronavirus, the refracted view is one in which the metaphorical eye focused primarily on physical health and the dangers posed by Covid-19; everything else was dispersed into the peripheral field of vision. When the phrase entered the new pandemic lexicon in the spring of 2020, the operative conception of

39 See generally 'Coronavirus (Covid-19): guidance and support' with links to various pages and guidance documents such as Public Health England, *Guidance for Managing a Funeral during the Coronavirus Pandemic* (first published in April 2020).

40 At least in Northern Ireland, Scotland and Wales where it appeared in the various regulations (see eg reg 12 of the Health Protection (Coronavirus Restrictions) (No 2) (Wales) Regulations (2020) (SI 2020/725 (W 162)) following an earlier, albeit more limited, directive in an amended version of the original principal regulations) but not in England where it retained a strongly recommended but nonetheless advisory status.

41 And this time across the four nations as government advice became a legal requirement in each one: see eg the Health Protection (Coronavirus, Wearing of Face Coverings in a Relevant Place) (England) Regulations 2020 (SI 2020/791).

42 This generic term denotes the combined statutory regulations and government guidance affecting funerals, unless it is necessary to distinguish between the two.

43 Repositories of the key changes, in each of the UK's four nations, can be found on the website of the Deceased Management Advisory Group (DMAG): see 'Government advice'. The DMAG was established early in the pandemic as the central co-ordination point for the funerals, bereavement and death care sector.

the 'serious and imminent threat to public health' mantra[44] became the respiratory illness caused by coronavirus and its immediate, observable impacts in confirmed cases, increased hospitalisations and virus-related deaths. As with all the forced changes to our lifestyles and deathstyles, it was this narrowly construed threat to public health that redefined Covid-era funerals. These became potential super-spreader events, with stories of localised clusters at burials and cremations confirming the risks involved.[45]

Utilitarianism as a moral framework for imposing swathes of legal restrictions that focused solely on the risk of death or serious illness to the living might seem like an obvious ethical choice, and one that was an essential crisis management tool when organised measures to stop the spread of coronavirus dominated the social and political landscapes. One could argue, however, that equating public health with physical health – and with preventing one disease to the exclusion of others – was an erroneous equivalence, with predictable outcomes. A myopic and temporally selective focus on Covid-19 at best downgraded, and at worst ignored, all other physical illnesses that presented during the pandemic.[46] Yet, this is only part of the picture. The conjoined elements of physical and mental wellbeing are central to any discourse around health as a basic human right and the measures that states must implement to ensure this.[47] While the psychological impact of coronavirus rules also assumed greater prominence as the pandemic progressed,[48] funerals – as events that are cloaked in emotionality – showcased the tensions between physical and mental health from the outset. In pursuing the legitimate goal of protecting the public as the dead became unwitting vectors for virus transmission, studies carried out by those working in the fields of bereavement and primary health

44 The justificatory basis for adoption of sweeping state powers from March 2020: see n 21 above and accompanying text.

45 See eg 'Coronavirus: seventeen family members get virus at funeral' *The Times* (London 30 March 2020). Similarly distressing stories emerged elsewhere: see eg 'Six people die from attending the same funeral in South Carolina' *The Independent* (London 17 April 2020).

46 The full impact of delayed diagnoses, rescheduled appointments and routine screening being missed will be felt for years to come. See eg C R Wells and A P Galvani, 'Impact of the Covid-19 pandemic on cancer incidence and mortality' (2022) 7 The Lancet Public Health e490-e491; R M Wood, 'Modelling the impact of Covid-19 on elective waiting times' (2022) 16 Journal of Simulation 101–109.

47 See eg the preamble to the 1946 Constitution of the World Health Organisation which defines health as a 'state of complete physical, mental and social well-being and not merely the absence of disease or infirmity'.

48 See eg B Pfefferbaum and C S North, 'Mental health and the Covid-19 pandemic' New England Journal of Medicine (13 April 2020); K Konstantinos, M Economou and C Papageorgiou, 'Mental health effects of Covid-19 pandemic: a review of clinical and psychological traits' (2020) 17 Psychiatry Investigation 491.

care are revealing the emotional impact of the legal rules, government guidance and resultant paradigm shifts in the funeral on those bereaved during the pandemic.[49] Before examining these impacts in more detail, it might be helpful to consider the socio-cultural function and significance of the funeral.

WHAT FUNERALS 'DO' AND THE IMPACT OF CORONAVIRUS RESTRICTIONS

Death is more than a biological progression; as Charmaz points out, 'it is an inherently social process'.[50] The absence of a clear dividing line between biological and social death has been highlighted by scholars across a range of disciplines,[51] and it is this duality that shapes our perceptions of the recently deceased who straddle the metaphysical boundary between person and thing. On the one hand, they are inanimate objects, devoid of physical life and undergoing the inevitable process of decay; at the same time, the newly dead remain socially alive, enmeshed within a network of relational ties that ensure the corpse is still narratively connected with the human being that was (and, in some ways, still 'is').[52] Funerals reflect this duality, by fusing the 'practical and emotional tasks associated with death'.[53] They are essential dispositive mechanisms that remove bodily matter from the active realm of the living, while cementing the deceased's place in the lives of family, friends and community through a series of socially mandated and highly symbolic acts. It is the latter component of the funeral – the set behaviours that frame the ritual in the minds of participants – that was fundamentally changed by Covid-19 and the associated process of law-making.

Funerals are major events that go beyond expressions of love and admiration for the deceased,[54] or the performance of religiously or

---

49 See the various sources listed throughout the fourth part below: 'The human cost of lockdown funerals: emerging evidence of a new public health crisis'.

50 K Charmaz, *The Social Reality of Death: Death in Contemporary America* (Addison-Wesley 1980).

51 See eg E Hallam, J L Hockey and G Howarth, *Beyond the Body: Death and Social Identity* (Routledge 1999); and K J Norlock, 'Real (and) imaginal relationships with the dead' (2017) 51 Journal of Value Inquiry 341.

52 See eg C Valentine, *Bereavement Narratives: Continuing Bonds in the Twenty-First Century* (Routledge 2008).

53 Bosley and Cook (n 17 above) 69.

54 In late modern Western societies, funerals are more about 'serving the needs of the bereaved, rather than commending the departed': M Holloway et al, 'Funerals aren't nice but it couldn't have been nicer: the makings of a good funeral' (2013) 18 Mortality 30, 30.

culturally mandated death customs.[55] Of course, there are dangers in essentialising, and implying that all funerals are experienced the same way: things like social class, ethnic origin, religion and geography have a major influence.[56] Yet, there are commonalities.[57] Funerals supply both a time and a normative framework for conveying sympathy to the bereaved,[58] who have social licence to display their emotions as part of the grieving process.[59] They also provide essential support for the bereaved,[60] creating a sense of order and stability at a time of chaos and disorder– what Norton and Gino describe as 'compensatory mechanism[s] designed to restore feelings of control after losses'.[61] And they offer societal acknowledgment of the change in relationships that death has triggered,[62] marking the deceased's permanent removal from the living community while 'assur[ing] survivors that the world goes on'.[63] Thus funerals, with their composite blend of the transformative, expressive and supportive, typically accomplish a range of social and psychological functions.

Unravelling the cognitive and affective underpinnings of rituals more generally, Hobson and others describe them as formalised, symbolic expressions that are slow to change, and whose meanings

55 These are essential constitutive elements of certain belief systems and cultures: see eg T O'Rourke, B H Spitzberg and A F Hannawa, 'The good funeral: toward and understanding of funeral participation and satisfaction' (2011) 35 Death Studies 729.

56 See eg T Walter, 'Three ways to arrange a funeral: mortuary variation in the modern west' (2015) 10 Mortality 173. For instance, geographically, practices will vary from country to country, but also within countries with local and regional variations and urban/rural divides.

57 And for a general discussion on funerary orthodoxy during epidemics, from which commonalities can be drawn, see S Ripoll, 'Death and funerary practices in the context of epidemics: upholding the rights of religious minorities' *CREID Working Paper 3* (Institute of Development Studies 2020).

58 V Lensing, 'Grief support: the role of funeral service' (2001) 6 Journal of Loss and Trauma 45; H B Mitima-Verloop, T T M Mooren and P A Boelen, 'Facilitating grief: an exploration of the function of funerals and rituals in relation to grief reactions' (2021) 45 Death Studies 735.

59 O'Rourke et al (n 55 above) 746. See also Bosley and Cook (n 17 above) 78.

60 W G Hoy, *Do Funerals Matter? The Purposes and Practices of Death Rituals in Global Perspective* (Routledge 2013).

61 M I Norton and F Gino, 'Rituals alleviate grieving for loved ones, lovers, and lotteries' (2014) 143 Journal of Experimental Psychology: General 266, 268 citing B D Romanoff, 'Rituals and the grieving process' (1998) 8 Death Studies 697.

62 Lensing (n 58 above) 49.

63 L A Gamino et al, 'Grief adjustment as influenced by funeral participation and occurrence of adverse funeral events' (2000) 41 OMEGA—Journal of Death and Dying 79, 79–80.

are reinforced through constant repetition.[64] Funerary customs are no different: they are 'long standing traditions and non-discursive practices' where things are done in a particular way and any deviation must be 'justified, explained, discussed, and negotiated'.[65] Rebay-Salisbury's depiction highlights the enduring quality of this particular ritual, and suggests a process of change that – when it occurs – is both deliberative and incremental. Recent transformations to funerals bear this out; the emergence of things like direct cremation (with no funeral service and no mourners), eco-friendly burials and personalised ceremonies,[66] has been gradual and prompted by a mix of ideological shifts, technological advances and contemporary social norms.[67] With Covid-19, however, funerals were reconfigured overnight and for very different reasons. Achieved by a potent and highly effective combination of legal rules and government guidance, the result was a set of swift, pervasive and highly disruptive changes to the embedded, ritualised content of all funerals through limits on attendance and the operationalisation of social-distancing measures.

'Stay at home' orders, numerical limits and a process of unnatural selection

Funerals are collective and participatory events that simultaneously create and reinforce specific death rites among those attending. The idealised narrative of a 'good funeral' speaks to a social grouping, to interactive and shared experiences where the presence of an 'audience … is crucial'[68] and gives the funeral its 'ritual potency'.[69] Covid restrictions, however, created a very different type of funeral medium.

At the start of the pandemic, the legal obligation to stay at home along with the imposition of numerical limits on funerals removed the community element, transforming the typical funeral from a

64 N M Hobson et al, 'The psychology of rituals: an integrative review and process-based framework' (2018) 22 Personality and Social Psychology Review 260, 260–261.

65 K Rebay-Salisbury, 'Inhumation and cremation: how burial practices are linked to beliefs' in M L S Sørensen and K Rebay-Salisbury, *Embodied Knowledge: Historical Perspectives on Technology and Belief* (Oxbow Books 2012) 15, 15. The funeral's 'formal standardisation and repeatability' are also noted in E Knopke, 'The arranged mourning ambience: about the professional production of atmospheres at funeral services' (2020) 25 Mortality 433, 434.

66 See eg Cook and Walter (n 9 above).

67 With the possible exception of direct cremation, the ritualistic elements of funerals remained intact: people still gathered together, engaging in acts of collective remembrance with some form of ceremony.

68 T Bailey and T Walter, 'Funerals against death' (2016) 21 Mortality 149, 152.

69 N Turner and G Caswell, 'A relative absence: exploring professional experiences of funerals without mourners' (2022) 85 OMEGA—Journal of Death and Dying 868, 871. See also O'Rourke et al (n 55 above).

public event that evokes a sense of 'symbolic communitas',[70] of connectivity and consequent emotional support, to an innately private and more isolated affair. For large parts of the population, such as rural communities, faith groups and ethnicities with traditions of large funerals, this change was a significant one.[71] Familiar, enduring patterns of 'gather[ing] at the same time in one place' and engaging in set behaviours as part of the funeral's implied 'ritual scripts'[72] became part of an alien landscape, replaced with strange new adaptive rituals such as lining the route to watch the funeral procession pass as a mark of respect for the dead and support for the living.[73] However, the cap on numbers also meant that funerals became 'invite only' events for the bereaved, as a maximum of what was originally only 10 members of the deceased's household or close family (or failing that, friends of the deceased) could physically attend.[74]

Walter and Bailey note that a funeral 'stratifies mourners into family or nonfamily',[75] but the changes imposed in March 2020 prompted a new process of intra-familial stratification, as surviving relatives were forced to decide which 10 individuals should have this particular 'golden ticket'.[76] Families have described making near impossible choices to whittle down attendance,[77] with potential for conflict due to

70 C Wouters, 'The quest for new rituals in dying and mourning: changes in the We–I balance' (2002) 8 Body and Society 1, 2.

71 The impact is outlined in *'A Good Death' during the Covid-19 Pandemic in the UK: A Report of Key Findings and Recommendations* (LSE Anthropology 2020) 10–14.

72 Knopke (n 65 above) 434.

73 Whether this simple act breached the prohibition on 'outdoor gatherings' (see eg reg 6A of the Health Protection (Coronavirus, Restrictions) Regulations (Northern Ireland) 2020 (SR 2020/55) as amended) was one of many interpretative questions posed by the measures. One option was to treat this as a series of individual gatherings, with no resultant breach: see D Holder, 'From special powers to regulating the lockdown: the Health Protection (Coronavirus, Restrictions) Regulations (Northern Ireland) 2020' (2020) 71 Northern Ireland Legal Quarterly 537, 549

74 However, some local authorities initially banned *all* mourners from their crematoria to limit the risk of virus-spread: see eg 'UK councils begin to ban funeral ceremonies due to coronavirus' *The Guardian* (London 4 April 2020) referring to the practice at Bradford, Leeds and Kirklees and 'Ban on mourners at Belfast Crematorium continues' *Belfast Telegraph* (Belfast 20 April 2020).

75 T Walter and T Bailey, 'How funerals accomplish family: findings from a mass-observation study' (2020) 82 OMEGA–Journal of Death and Dying 175, 175.

76 As Roald Dahl fans will know, the 'golden ticket' was the rare and coveted pass to enter Willie Wonka's Chocolate Factory, with only five tickets available worldwide: R Dahl, *Charlie and the Chocolate Factory* (1964).

77 See A Torrens-Burton et al, '"It was brutal. It still is": a qualitative analysis of the challenges of bereavement during the Covid-19 pandemic reported in two national surveys' (2022) 16 Palliative Care and Social Practice 26323524221092456.

restricted numbers.[78] The situation improved as permitted attendees increased in each of the four nations (rising for example, to 30 in England and 25 in Northern Ireland by October 2020),[79] and the legal restriction to members of the same household and close family was dropped.[80] Many families, however, had to endure a tough process of internal negotiation in finalising a short guest list for the deceased's funeral – especially during the critical first wave when restrictions were most severe. Virtual attendance, if the funeral could be live-streamed, became the default option for those who did not make the list.[81]

The concept of a hierarchy of mourners is not a new thing; within families, there is a sense in which funerals have always been about crafting a highly visible and perpetual marker of who was closest to the deceased in life.[82] For lockdown funerals, however, such internal rankings were conditioned by a mixture of external variables[83] and relational dynamics. No legal definition of 'close family member' appeared in the original statutory regulations, or in subsequent versions that adopted the same language.[84] This contrasts sharply with other categories of death laws, with their fixation on set and frequently hierarchical orderings of family that offer little in terms of

78 Funerals have always been stress amplifiers within families (see H Conway and J Stannard, 'The honours of Hades: death, emotion and the law of burial disputes' (2011) 34 University of New South Wales Law Journal 860). One suspects that decisions on who made the 'final cut' of 10 have created new fault lines within families.

79 At the same time, the legal limit in Scotland was 20, while numbers in Wales depended on venue size and availability of social distancing. Hickman et al (n 15 above) have criticised the numerous jurisdictional variances in the delegated legislation; with funerals, these would have created more confusion and distress for the bereaved.

80 This change occurred several months into the pandemic, under the second substantive set of regulations in each of the four nations: see eg the Health Protection (Coronavirus, Restrictions) (No 2) (England) Regulations 2020 (SI 2020/684) and the Health Protection (Coronavirus Restrictions) (No 2) (Wales) Regulations 2020 (SI 2020/ 725 (W 162)) which came into effect in early July 2020.

81 See eg S Pitsillides and J Wallace, 'Physically distant but socially connected: streaming, funerals, memorials and ritual design during Covid-19' in P Pentaris (ed), *Death, Grief and Loss in the Context of Covid-19* (Routledge 2021) 60–76.

82 Conway and Stannard (n 78 above).

83 Namely, practicalities and generic legal restraints where eg certain individuals could not attend because they were shielding, self-isolating or prohibited from travelling due to ongoing restrictions.

84 Likewise, there was no definition of 'a member of the deceased's household' (though this is easier to determine) or 'a friend' (more nebulous, by its very nature).

flexibility.[85] By omitting any definition or attempted categorisation, the regulations allowed the bereaved to construct and formalise their own definition of close family and to make specific choices at a time when so much freedom to choose other aspects of the funeral ritual (eg lowering the coffin into the grave, holding a wake) was denied to them. Research is emerging on whether the legal rules and the lived realities of a pandemic have reinforced normative concepts of kinship or constructed an alternative version of 'close family' in the funeral context.[86] However, the basic concept of funerals as 'must attend moments in the lives of all surviving family members'[87] was suddenly suspended by restrictions designed to tackle the pandemic.

Social distancing and new modes of funeral behaviour

Funeral rites have a normative force; they compel people to behave and to interact in certain ways, when participating in this universal post-death ritual. However, social distancing meant that those who could attend so-called 'Covid-safe' funerals had to maintain a two-metre distance from persons from a different household when indoors; this involved sitting and standing apart at all times in places of worship, crematoria and other funeral venues, in line with the prevailing legal requirements in Scotland, Wales and Northern Ireland.[88] Whether this also prohibited hugging or taking someone's hand, with consequential enforcement powers in the event that mourners literally reached out to one another, is open to interpretation.

We accept that the law restricts unwanted physical contact within both the criminal and civil law spheres,[89] yet to say that people were 'not allowed to hug' (an oft-cited mantra during the pandemic, and not just in the funeral context) or touch overstates the legal remit of social-distancing rules. Strict adherence to the rules in an enclosed

85 Eg s 47(a) of the Coroners and Justice Act 2009 lists 'interested persons' to any coronial investigation as including a spouse, civil partner, cohabiting partner, parent, child, sibling or step-parent of the deceased (in that order). S 65 of the Burial and Cremation (Scotland) Act 2016 allows the deceased's 'nearest relative' to arrange the funeral, defined as the surviving spouse or civil partner, then cohabiting partner followed by child or step-child, parent, sibling etc.

86 See eg J McCarthy, K Woodthorpe and K Almack, 'The aftermath of death as a family and community event: a sociological perspective on diverse responses and experiences' 15th International Conference on Death, Dying and Disposal (Manchester Metropolitan University 2021). Sociologists have already identified the funeral as a theoretical lens for analysing what constitutes 'family': see eg K Woodthorpe and H Rumble, 'Families and funerals: locating death as a relational issue' (2016) 67 British Journal of Sociology 242 and Walter and Bailey (n 75 above).

87 P Giblin and A Hug, 'The psychology of funeral rituals' (2006) 21 Liturgy 11, 18.

88 Though not in England: see n 40 above.

89 Offences against the person and trespass to the person are obvious examples.

space would have made such interactions physically impossible beyond members of the same household, and any resultant violation might not have been viewed as an egregious breach of social-distancing rules given the context. Outdoors, there was no such legal edict; maintaining a distance from others was public health advice that assumed a 'rule-like' status, but did not amount to a prohibition on hugging or other similar gestures.

Yet, in a pandemic that has highlighted how fundamental touch is to the human experience and the emotional consequences of 'touch hunger' when physical contact is stripped from everyday life,[90] the shared social ritual of the funeral illustrates the point perfectly. Virus suppression measures meant suppressing the innate behaviours and non-verbal communications that are embedded in the emotional fabric of this particular dispositive rite as acts of comfort and reassurance,[91] yet were suddenly classed as health-harming. For some, instinct trumped social-distancing rules; and when crematorium staff interrupted a service in Milton Keynes in October 2020, instructing two sons to move away from their distraught mother as they comforted her during their father's funeral, the family's reaction was one of outrage and disbelief.[92] There is no doubt that this was a heavy-handed enforcement of the rules.[93] However, the incident shows how a basic need to provide emotional support outweighed any risk of virus transmission *and* justified breaking the restrictions (at least in a moral sense). This latter point brings us to another important theme: that of compliance.

The compliance question

Transforming funerals into quick, minimalist and sparsely attended affairs that were literally devoid of the human touch would only contain virus spread in this particular social setting if restrictions were followed.

---

90 See eg J Durkin, D Jackson and K Usher, 'Touch in times of Covid-19: touch hunger hurts' (2021) 30 Journal of Clinical Nursing e4; R Visser, 'Losing touch? Older people and Covid-19' in Pentaris (n 81 above) 197–208.

91 Such interactional elements (eg the affirmation of emotion, gestures of comfort and expressing condolences) are part of what makes a 'good funeral': O'Rourke et al (n 55 above) 746.

92 A reaction shared, when footage of the incident circulated on social media: 'Funeral interrupted as grieving sons told to leave mother's side by crematorium staff' *The Telegraph* (London 5 October 2020). Many crematoria issued edicts around physical contact between mourners: 'Coronavirus: mourners attending funeral services asked to avoid hugging and hand-shaking' *The Independent* (London 19 March 2020).

93 A familiar theme during lockdown more generally, with oppressive policing of 'stay at home' orders in particular: see eg 'UK police chided for overzealous response to coronavirus lockdown' (*Reuters* 30 March 2020).

Two things suggest high levels of compliance. The first is the public's reaction to the so-called 'Partygate scandal', where funeral restrictions acted as a lightning rod for much of the opprobrium towards then Prime Minister Boris Johnson. As stories emerged of lockdown-breaking gatherings in Downing Street, one dominant counter-narrative was of grieving relatives who dutifully adhered to the rules, with drastically scaled-back burials and cremations.[94] The second, and perhaps more evidentially compelling, is the comparatively rare media reporting of rule-breaking at funerals and consequent criminal prosecutions.[95] Other clear and egregious breaches would almost certainly have been reported, given how the pandemic dominated news reports for months across television, radio and digital platforms.

In short, funeral restrictions seemed to be reluctantly accepted, as long as some attendance was possible,[96] and the rules were not applied with unfeeling bureaucracy[97] or openly flouted by others.[98] So what of the underlying reasons? Closer analysis suggests a number of explanations, though one thing is clear: to suggest that the law was the sole architect of these new societal norms overlooks a number of pandemic-centric influences and some discrete factors linked to the nature of the funeral itself.

Compliance theory posits a number of reasons why people obey the law, including legitimacy, moral obligation, heeding legal deterrents

---

94 See eg R Hadden, 'We obeyed Covid rules as our dad died. I'm angry the PM has dodged a partygate reckoning' *The Guardian* (London 20 May 2022). Boris Johnson was also heavily criticised by Conservative MP Aaron Bell, who obeyed Covid restrictions at a family funeral: 'Tory Backbencher asks if Boris Johnson "Thinks I'm a fool" for following Covid rules at grandmother's funeral' *The Independent* (London 21 January 2022).

95 Though isolated incidents did occur, invariably where the number of mourners exceeded the legal limit: see eg 'Covid Scotland: police were called to funerals due to large numbers gathering during lockdown' *The Scotsman* (Glasgow 4 October 2021); 'Covid-19: two £10,000 fines issued for 150-person funeral' (*BBC News Online* 22 January 2021). See also the examples at n 98 below.

96 Crematoria that banned mourners at the beginning of the pandemic were severely criticised for doing so: see n 74 above.

97 As in the Milton Keynes crematorium, noted above.

98 Media reports of public anger over a lack of censure for what were apparently clear breaches of coronavirus restrictions included the June 2020 funeral of IRA figure, Bobby Storey, attended by senior members of Sinn Féin who were part of an estimated crowd of 2,000 mourners ('Bobby Storey funeral: "different rules for different people" – angry reaction from people who buried loved ones during lockdown' *The Newsletter* (Belfast 31 March 2021) and the presence of over 150 members of the travelling community at a funeral in Kettering, Northamptonshire in November 2020 ('Covid-19: MP claims "outrage" at dropped charge for 150-guest funeral' (*BBC News Online* 23 April 2021)).

and (to a lesser extent) fear of punishment.[99] Cognitive scientists, meanwhile, argue that laws and their enforcement mechanisms are dependent 'on a broad consensus about the moral legitimacy of the rules' and 'what constitutes appropriate behaviour'.[100] There are shades of all this in adherence to lockdown restrictions in general, though it seems that possible criminal sanctions were not the main driver. Instead, compliance was shaped by a range of things, including substantive moral support for the rules, normative obligations to obey the law, social norms and lived environment.[101] Stannard refers to 'socially necessary rules' that were 'adhered to as much by consent as by coercion'[102] and posits a model of psychological jurisprudence based not just on the perceived practical legitimacy of the rules but on people's feelings about them and an 'internalized duty'[103] to behave in a certain manner. In other words, self-interest (not contracting Covid-19), the collective social responsibility to keep others – particularly the elderly and clinically vulnerable – safe (not spreading coronavirus), and ensuring that the National Health Service could still cope under pressure (safeguarding hospital capacity) elicited a consensual model of compliance.[104]

Running through all of this was something else that had not been widely experienced before: a pervasive feeling of social panic (especially during the first months of the pandemic), fuelled by the insidious and seemingly uncontrollable nature of this new infective agent.[105]

---

99 See eg T R Tyler, *Why People Obey the Law* (Princetown University Press 2006); A Licht, 'Social norms and the law: why people obey the law' (2008) 4 Review of Law and Economics 715.

100 E Fehr and U Fischbacher, 'Social norms and human cooperation' (2004) 8 Trends in Cognitive Science 185, 185.

101 E B Kooistra et al, 'Mitigating Covid-19 in a nationally representative UK sample: personal abilities and obligation to obey the law shape compliance with mitigation measures' (2020) Amsterdam Law School Research Paper (2020–2019). See also A Burton et al, 'Understanding barriers and facilitators to compliance with UK social distancing guidelines during the Covid-19 pandemic: a qualitative interview study' (2022) Behaviour Change 1, the authors noting that things like caring responsibilities, fatigue, emotional needs and constantly changing rules were barriers to compliance.

102 J Stannard, 'Engage, explain, encourage, enforce: therapeutic jurisprudence and the coronavirus lockdown' (2021) 25 European Journal of Current Legal Issues 2.

103 Ibid 11.

104 K Murphy et al, 'Why people comply with Covid-19 social distancing restrictions: self-interest or duty?' (2020) 53 Australian and New Zealand Journal of Criminology 477.

105 C A Harper et al, 'Functional fear predicts public health compliance in the Covid-19 pandemic' (2020) 19 International Journal of Mental Health and Addiction 1875.

The rapidly embedded 'threat to life' narrative acted as a centripetal force, vindicating emergency measures across the UK while ensuring widespread adherence to them.[106] Other elements were also used advantageously. In obscuring the distinction between public health guidance (where compliance is optional) and legal prohibitions (which attract criminal sanctions), Hickman has accused the UK Government of creating and exploiting a sense of 'normative ambiguity'.[107] This led to a 'powerful sui generis form of emergency regulatory intervention'[108] that drove compliance by leading people to believe that the legal requirements – and consequent limits on personal freedoms – were greater than they actually were.[109] Examples have already been noted in the funeral sphere, from social distancing in graveyards and cemeteries to refraining from hugging where confusion between the advisory and the mandatory helped segregate different households.

Finally, there are several funeral-specific factors that would have prompted compliance with the restrictions. First, funerals are carefully planned and sober affairs and, by their nature, are less open to breaches than unplanned parties or social gatherings. Secondly, the presence of 'middle' men and women– namely funeral directors and celebrants/officiants, as well as cemetery and crematoria managers – was also pivotal in advising families what was permitted, and discouraging potential rule-breaking.[110] Thirdly, the age profile of the dead and those mourning them was probably a factor: while Covid was fairly indiscriminate, mortality rates were higher in the over-70s,[111] and their partners, siblings and close friends (who were likely of a similar age profile) would have been particularly careful not to allow such funerals to be sites of virus spread. And, last but not least, is the existence of a different type of internalised duty (to borrow Stannard's

106 And, in this respect, was probably as good a deterrent against non-compliance as the threat of legal sanction.

107 T Hickman, 'The use and misuse of guidance during the UK's coronavirus lockdown' (SSRN September 2020). Also available via doi.

108 Ibid 2.

109 Empirical studies appear to confirm this: see eg N Finch et al, 'Undermining loyalty to legality? An empirical analysis of perceptions of "lockdown" law and guidance during Covid-19' (2022) 85 Modern Law Review 1419. Similar criticisms have been levied at governments elsewhere: see A Eustace, S Hamill and A Mulligan, *Public Health Law during the Covid-19 Pandemic in Ireland*, Public Policy Report of the Covid-19 Law and Human Rights Observatory (School of Law, Trinity College Dublin August 2021).

110 Funeral directors were instrumental here and risked fines where clear and blatant breaches of the rules occurred. One was fined £10,000 for not managing the event correctly where 150 people attended a funeral when the maximum number was 30: see *BBC News Online* (n 95 above).

111 R E Jordan, P Adab and K K Cheng, 'Covid-19: risk factors for severe disease and death' (2020) British Medical Journal 368.

term) when burying or cremating the dead. Within the complex of 'aesthetic and emotional services'[112] that make up the funeral, there is an overwhelming sense of bereaved families 'doing the right thing'. For Covid-era funerals, this meant obeying the restrictions and avoiding the resultant shame and social stigma had families acted otherwise.[113]

THE HUMAN COST OF LOCKDOWN FUNERALS: EMERGING EVIDENCE OF A NEW PUBLIC HEALTH CRISIS

As the pandemic ebbed and flowed, the rules around funerals were tweaked in tandem with the sequential relaxation and reinstatement of lockdown restrictions, before being largely removed through the spring and summer of 2021.[114] Yet, limited attendance and social distancing – the two things that catapulted us into a radically different type of dispositive ritual as the first wave of the pandemic gathered pace in March 2020 – remained firmly in place for long periods, as governments continued the delicate balancing act of allowing the living to mourn their dead while limiting the spread of Covid-19.

With a UK coronavirus death toll approaching 230,000 at the time of writing, and higher numbers of non-Covid deaths during the pandemic, millions of people endured burials and cremations where the usual rituals and social interactions were stripped away. So, returning to the earlier argument that fundamentally altered funeral arrangements were an overreach on public health grounds, what effect did the restrictions have on the mental health and emotional well-being of those bereaved during critical phases of the pandemic?

Some of the literature is still in its infancy, though discrete pockets of evidence are steadily emerging from short-term and longitudinal

112 Knopke (n 65 above) 435.

113 J Riley, 'The losses and affordances of pandemic restrictions on funerals' 15th International Conference on Death, Dying and Disposal (Manchester Metropolitan University 2021). Responding to a post-presentation question from the current author on families being conscious of legal rules around funerals, Riley suggested that bigger concerns were safety (not wanting to transmit coronavirus) and a sense of a funeral being an 'awful place [to be] told off' for breaking the rules. (Note of conversation on file with author.) Riley's paper was based on findings from an Economic and Social Research Council-funded project involving herself and other researchers at the University of Aberdeen: see 'Care in funerals' School of Divinity, History, Philosophy and Art History, University of Aberdeen.

114 While coronavirus restrictions were lifted in England, Scotland and Wales in July 2021 and in Northern Ireland in August 2021, some specific rules around funerals had already been eased. For example, numerical limits for funerals were lifted in England in May 2021, while Northern Ireland moved – from April 2021 – to permitted numbers based on a risk assessment of the relevant building or premises.

studies. At the outset, we must acknowledge that looking at the impact of funerals in isolation is simply not possible. The pandemic was an emotionally traumatic event that transformed people's lives, careers, education, familial and social interactions. Likewise, it is difficult to separate the impact of funeral restrictions from distressing experiences of illness and loss due to other aspects of pandemic management – for instance, people who died after becoming infected with coronavirus in formal health and care provision settings – and the likelihood of the restrictions compounding this.[115] However, the emotional trauma caused by lockdown funerals is documented in the literature as a grim legacy of the pandemic.

The initial coronavirus response obliterated long-standing social conventions around dying, this final exit from the living world that is the inevitable precursor to the funeral. Just like the idealised narrative of the good funeral, there is the idealised narrative of the so-called 'good death', where someone passes away peacefully and surrounded by family.[116] However, Covid-19 fatalities and other deaths in hospitals or care homes during the pandemic 'embod[ied] the attributes of a "bad death", making them particularly devastating for bereaved kin',[117] as visiting was banned or severely restricted to reduce the threat of virus spread. Dying alone[118] thus became another public health imperative. In these circumstances, some sort of redemptive funeral ritual was needed, to compensate for letting the deceased down at the end of life.[119] Sadly, the reality was very different as 'restricted funeral practices caused further upset' to grieving relatives.[120] Not having wider family and friends present at the funeral, travelling to and from funerals alone, maintaining physical distance throughout, and the 'fear of further contagion and … specter of further loss' [121] if coronavirus

115 These additional stressors are recognised in the relevant literature. For specific examples, see T R Jordan et al, 'The Covid-19 pandemic has changed dying and grief: will there be a surge of complicated grief?' (2022) 46 Death Studies 84; R A Neimeyer and S A Lee, 'Circumstances of the death and associated risk factors for severity and impairment of Covid-19 grief' (2022) 46 Death Studies 34.

116 See eg B Heart, P Sainsbury and S Short, 'Whose dying: a sociological critique of the "good death"' (1998) 3 Mortality 65.

117 D Carr, K Boerner and S Moorman, 'Bereavement in the time of coronavirus: unprecedented challenges demand novel interventions' (2020) 32 Journal of Aging and Social Policy 425, 425.

118 In the sense of having no family (or friends) physically present.

119 Torrens-Burton et al (n 77 above).

120 Ibid 9.

121 Neimeyer and Lee (n 115 above) 39.

infections were to occur at the burial or cremation, all contributed to a sense of not giving the deceased the funeral that they deserved.[122]

The longer-term impact on the grieving process is also becoming clear. Funerals simultaneously normalise and facilitate this process;[123] they create a 'sense of being connected to a large community',[124] providing space to celebrate the individual's life. As such they are one-off, immensely therapeutic events with a strong psycho-social element, and impossible to replicate. All of this contrasts sharply with lockdown funerals, and the absence of the traditional wakes and funeral gatherings that bookend this particular dispositive ritual.[125] Experts agree that Covid-driven restrictions have triggered an increase in complicated grief disorders, where intense feelings of loss are prolonged and impair the individual's ability to function.[126] Enforced isolation and separation are dominant themes. Not only were the living isolated from the dying at the point of death; they were isolated from close relatives at the funeral, and deprived of the basic human need to interact physically with each other – and a community of mourners – following the deceased's passing.[127] Beyond the funeral itself, lockdown restrictions reinforced feelings of loneliness and social

122 Torrens-Burton et al (n 77 above) 9. Giving the deceased the 'send-off' they deserved is important for the bereaveds' sense of 'funeral satisfaction': see J Rugg and S Jones, *Funeral Experts by Experience: What Matters to Them Research Report* (University of York 2019).

123 Giblin and Hug (n 87 above) 11.

124 Wouters (n 70 above) 2.

125 Wakes and funeral gatherings were prohibited by restrictions on gatherings, rules on non-mixing of households and the closure of hospitality for long periods.

126 See eg M Stroebe and H Schut, 'Bereavement in times of Covid-19: a review and theoretical framework' (2021) 82 OMEGA—Journal of Death and Dying 500; S Albuquerque, A M Teixeira and J C Rocha, 'Covid-19 and disenfranchised grief' (2021) 12 Frontiers in Psychiatry 638874; F Diolaiuti et al, 'Impact and consequences of Covid-19 pandemic on complicated grief and persistent complicated bereavement disorder' (2021) 300 Psychiatry Research 113916; and K Doka, 'Grief in the Covid-19 pandemic' in Pentaris (n 81) 29, 30–31. See also Torrens-Burton et al (n 77 above); Neimeyer and Lee (n 115 above); Jordan et al (n 115 above); and the various other sources cited immediately below.

127 Substitute networks through eg live-streaming of funerals and virtual wakes played an important role. Yet some users found these to be inauthentic and lacking in emotional connectivity; as feeling 'voyeuristic'; and being problematic due to technology glitches: D Rawlings, L Miller-Lewis and Jennifer Tieman, 'Impact of the Covid-19 pandemic on funerals: experiences of participants in the 2020 Dying2Learn Massive Open Online Course' (2022) OMEGA—Journal of Death and Dying.

isolation and were akin to a 'second bereavement'.[128] Adhering to the rules meant that the bereaved were sequestered away, shut off from the outside world and from their usual support networks (both familial and social). Grieving alone, while coping with a heightened sense of anxiety around the pandemic and constant daily media coverage made 'grief and trauma [feel] inescapable',[129] exacerbating psychological problems caused by the deceased's passing and subsequent funeral.

Of course, the evidence that is emerging merits some measure of caution, given that more attention is likely to focus on negative experiences of lockdown funerals.[130] However, several positives have been identified. For some, any form of ceremony was appreciated because 'it was the best way possible at that time'.[131] Small(er) gatherings were viewed as more 'intimate' and 'personal' events[132] that were less pressurised.[133] Meanwhile, live-streaming fostered connectivity by enabling those who could not be physically present to experience the funeral (albeit in an altered medium).[134] What mattered was creating a meaningful occasion, not large-scale funeral participation.[135] Yet, the bulk of the studies referenced here depict something very different. For many people bereaved during the pandemic, normal grief experiences were disrupted by restrictions on funerals alongside the 'multiple ambiguous losses embedded in restrictions to everyday life'.[136] The result of this combination of law and government guidance is being described as a different type of public health crisis for large parts of the population, linked

128 C Pearce et al, '"A silent epidemic of grief": a survey of bereavement care provision in the UK and Ireland during the Covid-19 pandemic' (2021) 11 British Medical Journal Open e046872.

129 Torrens-Burton et al (n 77 above) 9-10.

130 And surveys are more likely to recruit participants with negative experiences.

131 H B Mitima-Verloop et al, 'Restricted mourning: impact of the Covid-19 pandemic on funeral services, grief rituals, and prolonged grief symptoms' (2022) 13 Frontiers in Psychiatry 1103.

132 LSE Anthropology (n 71 above) 7.

133 They relieved the pressure of having to host, or perform for, lots of people: Riley (n 113 above).

134 Pitsillides and Wallace (n 81 above), though not all virtual experiences were positive: see n 127 above.

135 A Burrell and L E Selman, 'How do funeral practices impact bereaved relatives' mental health, grief and bereavement? A mixed methods review with implications for Covid-19' (2022) 85 OMEGA—Journal of Death and Dying 345.

136 H Kaur, A K Lillie and C Wagstaff, 'Prognosticating Covid therapeutic responses: ambiguous loss and disenfranchised grief' (2022) Frontiers in Public Health 3.

to complicated grief.[137] And the effects, for those individuals, will continue to be felt for a long time.

CONCLUSION

> The usual rituals, customs and interactions that occur in the context of end of life and after a death are another casualty of the virus.[138]

As part of the sweeping changes imposed in early 2020, Covid-19 necessitated a temporary deprivation of certain rights surrounding death and mourning. When we think of the embedded narrative of the recently deceased as potential contaminants, invoking public health measures to mitigate disease spread in dealing with the dead is nothing new. However, the nature, scope and reach of the restrictions was unprecedented,[139] and used to target a very different threat in the funeral context: that of the living congregating around the corpse.

Covid-era funerals are prime examples of what Chua and Lee describe as 'governing through contagion'[140] with an amalgam of legal controls and public health directives deployed to ensure that private citizens did not become potentially lethal infective agents. As with most pandemic-induced measures, one could argue that the public health outcomes legitimised the means when focusing solely on virus transmission. With funerals, however, the human consequences stretch far beyond coronavirus risks. For the millions bereaved during the pandemic, grief is even more traumatic and unresolved when set against the backdrop of social isolation and the lived experience of funerals where the usual participatory elements, communal rites and support mechanisms were missing. There is a strong argument for saying that the Government got it wrong here, and that its focus on protecting the population against Covid-19 was overly narrow because it ignored the wider and equally important issue of mental health.[141] To date, most of the studies carried out on funerals support this point.

---

137 Pearce et al (n 128 above). This was predicted early in the pandemic: see eg S S Mortazavi et al, 'Fear, loss, social isolation, and incomplete grief due to Covid-19: a recipe for a psychiatric pandemic' (2020) 11(3.Covid19) Basic and Clinical Neuroscience 225.

138 Albuquerque et al (n 126 above) 2.

139 Probably one of the most over-used words of the pandemic.

140 L Chua and J J G Lee, 'Governing through contagion' in V V Ramraj (ed), *Covid-19 in Asia: Law and Policy Contexts* (Oxford University Press 2021) 116–132.

141 Especially at the critical phases of the pandemic when the restrictions were harshest, though regrets were later expressed about this: 'Banning close relatives from funerals during lockdown was wrong, admits Matt Hancock' *The Telegraph* (London 10 July 2020).

By drawing on research from other disciplines to gain a deep and informed understanding of the wider impact of the restrictions, this article has highlighted the dynamic effects of law and public health protocols on a highly symbolic social ritual. Like so many of the rules that were rushed through during the pandemic, the longer-term effects on individual and collective experiences of bereavement, and their consequent emotional toll, are only being understood with the passage of time. Reflecting on these findings not only enables us to identify the strengths and weakness of the coronavirus measures. It also alerts governments to the need to frame better solutions when the next pandemic or similarly catastrophic event strikes – solutions that are not just short-term, emergency management processes, but drafted with potentially adverse outcomes in mind.

Finally, there is the question of lasting effects on the funeral itself, and whether pre-pandemic levels of attendance and behavioural norms will be restored post-pandemic. It may be that this final rite of passage has been permanently transformed, and that coronavirus has simply accelerated what some sociologists had already identified as the diversification of funerals norms alongside the 'privatisation' of contemporary funerals as smaller, invite-only events.[142] Yet, such claims may be premature. The 'traditional' funeral,[143] with its public and communal dimensions, its sense of bringing people together to emote and to offer support at a time of profound loss, is embedded in our socio-cultural DNA.[144] Short-term reversals are not accurate bellwethers of long-term trends, especially in a pandemic where changes were imposed by the state and not driven by personal choice. The funeral ritual will remain, even if its form alters. Only time will reveal whether the final rite of passage, as a collective experience with a physically present and large(r) supporting cast, is something that people re-embrace on emerging from living and dying in the shadow of coronavirus.

142 K Woodthorpe et al, '"My memories of the time we had together are more important": direct cremation and the privatisation of UK funerals' (2022) 56 Sociology 556.

143 Which is an emotionally laden term in itself, and infers a certain resistance to change.

144 And especially so, for specific religions and cultures – and for certain generations of people.

Northern Ireland
Legal Quarterly
Summer Vol. 74.2 (2023) 237–268
Articles DOI: 10.53386/nilq.v74i2.1034

Journalistic privilege in Ireland

Cian Ó Concubhair*
Maynooth University

Correspondence email: cian.oconcubhair@mu.ie

ABSTRACT

Legal protection for confidential journalist sources has often been a site of tension and dispute between journalists, the police and the courts. Journalists routinely claim that freedom of expression guarantees provided for under international and domestic human rights instruments include a legal privilege against disclosure of confidential journalist sources. This claim is often raised to resist compelled disclosure of journalistic materials to police as part of criminal investigations. Courts in many jurisdictions have forcefully repudiated this legal claim, though many recognise some right for journalists to refuse disclosure. Some courts have reluctantly conceded to the naming of this right as 'journalistic privilege'.

In 2020, courts on both sides of the Irish border were called upon to vindicate this right against disclosure. This recent flurry of litigation has, in the Republic of Ireland, built upon more than a decade of significant legal developments around 'journalistic privilege'. These latter developments have dramatically expanded the scope of the Irish Constitution's freedom of expression guarantees.

This article critically reviews this last decade of significant legal developments around 'journalistic privilege' in the Republic of Ireland. It examines the two recent and highly significant Irish determinations from 2020 in *Fine Point Films* and *Corcoran*, and how the former Northern Irish judgment has created significant new avenues for legal development in the Republic of Ireland. The article also identifies and considers some important, emergent themes in Strasbourg's article 10 jurisprudence: specifically an apparent new 'source motive' test for article 10 protection of confidential source material.

Keywords: journalistic privilege; confidential source protection; article 10 ECHR; article 40.6.1º Bunreacht na hÉireann; media freedom; police powers.

* I would like to thank the NILQ's anonymous reviewers for their feedback and comments on the initial submission of this article. I would also like to thank Ronan Lupton SC for reviewing an early draft of this piece.

INTRODUCTION

In 2020, appeals courts north and south of the Irish border were called upon to consider the question and scope of 'journalist privilege' in two strikingly similar cases. In *Fine Point Films*[1] and *Corcoran*[2] the High Courts of Northern Ireland and the Republic of Ireland (hereinafter 'Ireland') were each faced with journalistic challenges to police seizure of property during court-authorised searches of their workplaces and homes. Both cases drew heavily upon the article 10 jurisprudence of the European Court of Human Rights (ECtHR) regarding confidential source protection and adapted and applied it within the specific procedural regimes governing court authorisation of police searches. *Corcoran* also added significantly to the evolving judicial debate around the scope of journalist source protection under article 40.6.1º, Bunreacht na hÉireann 1937: Ireland's principal human rights instrument.

Yet, despite the many similarities between the cases, these courts came to quite different conclusions. The Northern Ireland High Court in *Fine Point Films* offered a full-throated endorsement of the need for an *overwhelming public interest* to pierce the European Convention on Human Rights (ECHR) article 10 corollary *right* of confidential source protection.[3] By contrast in *Corcoran*, whose hearings immediately followed and considered *Fine Point Films*, the High Court in Ireland saw judicial ambivalence towards an assertion of 'journalist privilege', continuing a long and consistent trend in Irish (and other common law)[4] courts.

Corcoran was, however, appealed to the Irish Court of Appeal in 2022. In this subsequent *Corcoran*[5] decision, the Irish Court of Appeal overturned the High Court judgment and adopted the approach of the Northern Irish court in *Fine Point Films*. Like *Fine Point Films*, the Court of Appeal judgment sets out a number of important principles governing 'journalistic privilege' in Ireland. The judgment also fired a significant warning shot to the Oireachtas regarding the potential for findings of unconstitutionality against aspects of the Irish warrant-granting regime. In June 2023, the Irish Supreme Court unanimously

1 *In Re Fine Point Films & Others* [2020] NICA 35.

2 *Corcoran v An Garda Síochána* [2020] IEHC 382.

3 See eg *Goodwin v United Kingdom* (1996) 22 EHRR 123 and *Financial Times Ltd v United Kingdom* (821/03) [2010] EMLR 21.

4 Janice Brabyn, 'Protection against judicially compelled disclosure of the identity of news gatherers' confidential sources in common law jurisdictions' (2006) 69 Modern Law Review 895; and Jeffrey Nestler, 'The underprivileged profession: the case for Supreme Court recognition of the journalist's privilege' (2005–2006) 154 University of Pennsylvania Law Review 201.

5 *Corcoran v An Garda Síochána* [2022] IECA 96.

endorsed the Court of Appeal's judgment in *Corcoran*, and rejected an appeal by An Garda Síochána.[6]

These recent Irish judgments raise a number of distinct, yet interrelated, complex issues for both Irish constitutional law and European (and Northern Irish) human rights law. The article begins by examining the significant developments around 'journalist privilege' in Ireland, where the issue of confidential source protection has been vented and explored by appellate courts on a number of occasions during the past decade and a half. It then considers the High Courts' judgments in *Fine Point Films* and *Corcoran* and their implications for journalist privilege on both sides of the Irish border, with a particular emphasis on the latter's novel development of Irish constitutional protections for publishers. The article then sets out and critically reviews the most recent of the three cases examined here, the Court of Appeal judgment in *Corcoran*. The article concludes by outlining and analysing some shared emergent themes in journalist source protection on the island of Ireland and sets out the key legal principles that appear to govern both legal regimes.

This article provides a comprehensive critical re-evaluation of the history of journalistic privilege in Ireland at a highly significant moment in its development. The article also provides a number of novel theoretical insights into 'journalistic privilege' in Ireland. In particular, the article examines and theorises a decades-long tension between Ireland's judicial branch and its news media regarding which institution sits at the apex of constitutional importance in safeguarding Irish democracy. This issue has, I demonstrate, played out most explicitly in cases where claims of 'journalistic privilege' were at issue. It also draws attention to related emergent trends in the European Court of Human Rights' article 10 jurisprudence, particularly what I describe as a new 'source motive' test for confidential source protection. This article strongly argues this previously critically unexamined line of Strasbourg authority may have significant implications for crime-reporting journalism in Europe.

6 *Corcoran v An Garda Síochána* [2023] IESC 15.

DEVELOPMENT OF 'JOURNALIST PRIVILEGE' IN THE REPUBLIC OF IRELAND: 'WHO DECIDES' AS A DEMOCRATIC CONSTITUTIONAL IMPERATIVE

No special status for news media: article 40.6.1º and the scope of the 'administration of justice'

While the Irish Constitution expressly provides liberty of expression for journalists and news media under article 40.6.1º,[7] journalists seeking a corollary right of source confidentiality protection (*qua* 'privilege') have had mixed treatment in the Irish courts. Beginning in *Re O'Kelly*,[8] the Court of Criminal Appeal delivered a robust 'administration of justice in the courts'[9]-centred interpretation of article 40.6.1º that dismissed the defendant's claims that it might afford journalists a distinct set of protections. While article 40.6.1º might protect the 'education of public opinion' work of journalists because of that work's contribution to the 'common good', Walsh J determined these protections did not extend to any special right of non-disclosure of source materials, let alone the seeming absolute privilege asserted by various journalists' codes of ethics.[10]

> So far as the administration of justice is concerned the public has a right to every man's evidence except for those persons protected by a constitutional or other established and recognised privilege.[11]

The court in *Re O'Kelly* concluded that there was no 'established and recognised' privilege for journalists in Irish law. Journalists, like any other ordinary citizen, were obligated to disclose information to a court that required it. Not only was there no 'privilege' against disclosure, journalists could not expect any additional leeway from the courts despite the asserted necessity of confidential source networks in their work.

This reference to the 'administration of justice' in *Re O'Kelly* hints at the expansive borders of judicial power the Irish courts have carved out under article 34.1. *Re O'Kelly* was also decided soon after *Murphy v Dublin Corporation*.[12] While not a 'journalist privilege'

---

7 Carolan notes the unique pedagogic function the provision envisages for news media. Eoin Carolan, 'The implications of media fragmentation and contemporary democratic discourse for "journalistic privilege" and the protection of sources' (2013) 49 Irish Jurist 182.

8 (1974) 108 ILTR 97.

9 See art 34.1, Bunreacht na hÉireann 1937.

10 See eg the National Union of Journalists, 'Code of Conduct'.

11 *Re O'Kelly* (n 8 above).

12 [1972] IR 215. This connection was identified by Simons J in *Corcoran* [2022] (n 5 above) [35].

case, the Supreme Court in *Murphy* had similarly 'emphasised that the decision as to the compellability of the production of evidence is a matter for the *judicial power*'.[13] These ruminations – repeated again in later Irish judgments – on the proper sphere of judicial authority when repudiating the asserted 'journalist privilege' might lead some to believe that this is a dispute within the Irish Constitution's separation of powers doctrine.[14] This would certainly be a novel approach: one that might afford journalism an enhanced constitutional status, on a par with the executive, legislative and judicial branches of government. Indeed, such a position may have some textual support. Article 40.6.1º itself refers to various news media as 'organs': in this case 'organs of public opinion'. This is a distinctive label in the 1937 Constitution, only afforded to the three recognised branches of government, institutions representing the state in international affairs and the media.[15] This novel conceptualisation of a freedom of expression instrument is no doubt inspired by the general corporatist influences in the Irish Constitution.[16] The corporatist dimension to article 40.6.1º might also support an interpretation that the news media as an institution – rather than more liberal individualist conceptions of freedom of expression – had some special constitutional status. Despite these apparent textual possibilities, however, the Irish courts have generally been unenthusiastic in their development of article 40.6.1º.[17] Until very recently, article 10 ECHR was treated as offering significantly more protection to journalists in Ireland.[18] If the framers of the Irish Constitution had possibly sought to recognise news media's institutional power within that document, the Irish courts nullified that ambition.

Instead of identifying and exploring such textual, institutional and democratic possibilities, the court in *Re O'Kelly* was instead anxious to assert the superior decision-making authority of the judiciary – over journalists – in determining whether source confidentiality should be protected. *Re O'Kelly*'s forthright judicial repudiation of the existence of a special form of protection for journalists' sources was subsequently

13 Ibid [35] (emphasis added).

14 See G Hogan, G Whyte, D Kenny and R Walsh, *Kelly: The Irish Constitution* 5th edn (Bloomsbury 2018) [3.2.112].

15 See arts 6.2, 29, 39, Bunreacht na hÉireann 1937.

16 For some discussion of the corporatist influences in Bunreacht na hÉireann 1937, see D K Coffey, *Drafting the Irish Constitution 1937–37: Transnational Influences in Interwar Europe* (Palgrave 2018).

17 See Hogan et al (n 14 above) [7.6.07] and *The Report of the Constitution Review Group* (Pn 2632, Stationery Office 1996) 292.

18 For example, the Supreme Court in *Mahon v Keena* [2010] 1 IR 336 relied exclusively on art 10 ECHR in recognising some form of 'journalistic privilege', without any reference to the textually generous art 40.6.1.

affirmed by Finlay CJ in the Supreme Court judgment *Burke v Central Independent plc*.[19] This early assertion of judicial authority is part of what I label the 'who decides' question: 'who decides' if there is to be any protection against disclosure, the journalist asserting privilege, or the courts? This 'who decides' question continued into the seminal Supreme Court judgment of *Mahon v Keena*, where some form of 'journalistic privilege' was first recognised in Irish law.[20] This question has, I suggest, become a defining feature of the law on journalistic privilege in Ireland.

It is worth reflecting for a moment on this judicial territorialism in relation to the 'who decides' question. Irish courts are not alone in their scepticism towards assertions of 'journalistic privilege'. Hostility towards perceived attempts by journalists to rob the courts of their authority to decide whose evidence is heard is part of a wider trend of judicial ambivalence across the common law world.[21] While recognising this international trend, the Irish Constitution may, however, provide some additional textual fuel for the antagonism of Irish courts. Article 34.1 does, after all, contain a generously broad, and rigidly interpreted,[22] requirement that '[j]ustice shall be administered in courts'.[23]

Through the Irish courts' article 34.1 lens, then, the 'who decides' question is elevated from an evidential principle to an imperative to defend the constitutional order. On this view, 'who decides' is a matter for the 'judicial power', not the media 'organ'. Recalcitrant assertions of privilege by journalists challenge the constitutional authority, and status, of the courts in Irish democracy.

Perhaps curiously, despite multiple references to the 'judicial power' and 'administration of justice' in the early Irish decisions on journalistic privilege, the courts have not explicitly cited article 34.1. This is a trend that has continued into the more recent judgments examined in this article. An obvious response might be that the separation of powers doctrine only includes the three branches of government, which by definition excludes the so-called 'fourth estate' of the media. However, the courts have used article 34.1 to repudiate other forms of non-state power.[24] Another reason might be that the Irish courts wish to avoid, despite this strong textual basis, ascribing any enhanced institutional

19 [1994] 2 IR 61.
20 *Mahon v Keena* (n 18 above).
21 Brabyn (n 4 above).
22 See Hogan et al (n 14 above) [4.2.12]–[4.2.15].
23 Bunreacht na hÉireann 1937, art 34.1.
24 See eg *Law Society of Ireland v Malocco* [2005] IESC 5.

status for the news media in Ireland's constitutional order, though this cuts against routine judicial acclamations of the central importance of the news media to the proper functioning of democracy.

Whatever the reason, this potential sensitivity around recognising or ascribing an enhanced status to news media has manifested itself in other, perhaps unfortunate, ways. For example, in its 1996 *Report on Contempt*,[25] the Irish Law Reform Commission took the opportunity to consider whether Ireland should adopt a legislative measure similar to section 10 of the Contempt of Court Act 1981 in the United Kingdom (UK). This gives a presumptive right to a journalist (or any other 'publisher') to refuse to disclose confidential sources, unless 'the interests of justice, or national security, or for the prevention of disorder or crime' require it.[26] While not the absolute privilege against disclosure asserted by journalists, section 10 does recognise an enhanced degree of autonomy for news media to resist judicial demands for evidence. It also, importantly, still preserves the ultimate authority of the courts on the 'who decides' question: a court can after all decide to pierce the 'privilege' if the listed interests are in play. A minority of the Commission endorsed the proposed change, but one which would have included a threshold of 'necessity' for judicial encroachment to further enhance the protection of journalist sources.[27] The Commission's majority rejected the proposal, as they were evidently concerned by the potential impingement on existing judicial authority.[28] This rigid refusal to concede any ground to the status of news media, may, as we shall see below, have left Irish law in constitutionally murky waters.

Mahon v Keena: recognition and repudiation

Judicial antipathy towards claims of journalistic exceptionalism enjoyed what appeared to be a substantial reversal in the Supreme Court's seminal judgment in *Mahon v Keena*. Here, for the first time, the Irish Supreme Court endorsed the validity of the *Goodwin*-line of article 10 ECHR in Irish law. In contrast to Irish courts until this point, the ECtHR in *Goodwin* offered an enthusiastic account of the centrality of journalistic autonomy under article 10 in realising and maintaining a healthy democracy. The media's recognised democratic value also

25 Law Reform Commission, *Report on Contempt of Court* (LRC 47–1994).

26 Section 10 has itself been criticised for not affording appropriate value weighting to the journalistic interest in source protection, and for containing exceptions that have been interpreted too expansively. See Brabyn (n 4 above) 916–917.

27 Law Reform Commission (n 25 above) 21.

28 Ibid 22. For an excellent discussion of the scope of judicial sensitivity around encroachment into their sphere of 'administration of justice', see Law Reform Commission, *Issues Paper: Contempt of Court and Other Offences and Torts involving the Administration of Justic*e (LRC IP 10–2016).

required, as a corollary, a general right to resist disclosure of confidential source information. The rationale of the ECtHR in *Goodwin* was that, to be an effective accountability tool in a democracy, journalists needed to maintain trust with their confidential sources. Per-*Goodwin*, article 10 requires the courts to respect the integrity of information-gathering networks. This allows for a presumptive journalist 'privilege' against disclosure, which required 'an overriding requirement in the public interest'[29] before disclosure should be ordered by the courts. Notably, for our purposes, *Goodwin* is unconcerned with the 'who decides' question: that judgment presupposes that courts remain the ultimate determiners of the application of article 10.

The 'who decides' question was, however, central to part of the outcome in *Mahon v Keena*. In that case, the defendant journalist Colm Keena had, on the direction of his editor, destroyed the relevant source materials when it became clear proceedings to identify his confidential source were imminent. Drawing on the experience of journalists in neighbouring Britain, his editor determined she could not trust the Irish courts to respect the practical need for source confidentiality.[30] While the court may have surprised the *Irish Times* editor when it found in the newspaper's favour, in its application and incorporation of *Goodwin*, the Supreme Court was at pains to reiterate the continuing supremacy of judicial authority in determining whether this 'privilege' was indeed pierced, and disclosure required.[31] The *Goodwin* version of 'journalist privilege', like article 10 from which it is derived, is not absolute, and it is for the courts to decide whether 'an overriding public interest' necessitates overriding the protection. For hijacking that undisturbed (by article 10) judicial role in answering the 'who decides' question, the defendant newspaper was duly punished by the Supreme Court. While the newspaper won the main legal argument around article 10 ECHR's applicability, the court made an unprecedented costs order against the victorious newspaper. The newspaper's subsequent challenge to the costs order's alleged hollowing-out of its 'privilege' in Strasbourg failed.[32]

The prospects for success in that ensuing appeal to Strasbourg were, after *Sanoma Uitgevers v Netherlands*,[33] likely to be modest. Though *Sanoma* similarly involved an article 10 ECHR success for journalists seeking to protect confidential sources from criminal investigations by police, the Grand Chamber's decision ultimately

29 *Goodwin* (n 3 above) [39].
30 Geraldine Kennedy, 'A cold, calculated decision to step outside the law' *Irish Times* (Dublin 25 October 2014).
31 *Mahon v Keena* (n 18 above) [92].
32 *Keena v Irelan*d (2014) 29804/10.
33 [2010] (Application no 38224/03).

revolved around the absence of an independent judge in authorising and supervising the search and seizure of journalistic materials. The failure of the Netherlands' particular warrant-granting procedures to require a supervising judge is what generated the article 10 breach. If an independent judge had authorised it, any search and seizure would, the Grand Chamber determined, have been permissible under the Convention.

On the 'who decides' question then, the Convention is clear: the courts must decide. While article 10 ECHR may celebrate the central role of the news media in our democracies. In Strasbourg's view then it is the courts, not journalists, that are the ultimate institutional guarantors of democracy.

Constitutionalising journalistic privilege

The last significant decision of note here is Hogan J's novel and refreshing reconsideration of the scope of journalistic privilege under article 40.6.1º Bunreacht na hÉireann in *Cornec v Morrice*.[34] While noting the then recent Supreme Court judgment in *Mahon v Keena* which focused exclusively on article 10 ECHR, Hogan J instead opted for a practically advantageous[35] route of recognising confidential source protection as also being a part of the Irish Constitution's fundamental rights. This was the first time such a journalistic right or interest under the Constitution had been observed by an Irish court. Hogan J's judgment in *Cornec* provides an enthusiastic re-evaluation of the express wording of article 40.6.1º, determining it gives at least equal – if not superior – protections as the article 10 ECHR.[36] *Cornec* emphasised the importance placed by the Irish Constitution on both the 'democratic nature of the state',[37] and the educative value in free information flow and dissemination:[38] 'essential in a free society'.[39] While 'journalists are central to that entire process',[40] anyone engaging in comparable activities (eg bloggers,[41] other 'citizen journalists', and we might assume, research academics) enjoys a degree of enhanced recognition and protection by the Constitution.

---

34 [2012] IEHC 376.

35 See Carolan (n 7 above).

36 An approach long-advocated by some commentators: see eg Eoin O'Dell, 'So, does Irish law now recognise a journalist source privilege?' (*cearta.ie* 9 August 2009).

37 Bunreacht na hÉireann, art 5.

38 Ibid art 40.6.1.

39 *Cornec* (n 34 above) [46]. For an excellent discussion on this pivot by the courts in constitutionally recognising an enhanced role for journalists in Irish democracy, see Carolan (n 7 above).

40 *Cornec* (n 34 above) [42].

41 Ibid [66].

Thus, in at least one important respect, *Cornec* echoed Walsh J's and other earlier Irish courts' rejection of a specific confidential source 'privilege' limited to 'bona fide'[42] journalists.[43]

According to Hogan J, any decision by a court to pierce this general 'privilege' should balance the express high constitutional value placed on public information dissemination and related democratic debate against the competing interest in disclosure. Interestingly, the court in *Cornec* also found that the public interest in affording protection to journalists under article 40.6.1º provided presumptive confidentiality to both the identity of sources *and* the information they imparted to the journalist.[44] The constitutional 'privilege' is not merely about source identity protection. For Hogan J, the corollary article 40.6.1º right attached to that which *the source* wished to be protected. While this focus on the particular source's interest in what should be kept confidential – identity or other information – is more expansive than article 10 ECHR, it is also consistent with the overarching objective of the privilege to generally maintain trust in the journalist–source relationship, and support the integrity of journalist information-gathering networks.

Given the conservative attitude the Irish courts have tended to show towards engaging with the express recognition of the special status of journalists under article 40.6.1º of the Constitution[45] – and the mixed blessings of *Mahon v Keena* – Hogan J's richly complementary account of the constitutional role of the media in Irish democracy raised the potential for a change in judicial–media relations. *Cornec* appeared to offer some recognition of the parity of esteem of news media and journalists in safeguarding and enriching Irish democracy:[46] something preceding judgments, with their obsessive reassertion of judicial supremacy in the constitutional democratic order, seemed disinclined in conceding. However, as with *Goodwin*, it is worth noting that the 'who decides' question was not engaged here: indeed, the court's consideration in *Cornec* of the threshold for judicial overriding of protections again presupposes judicial supremacy.

The principles of journalist privilege in Ireland following *Cornec* can be summarised thus:

---

42 See 'New practice direction bans "hobby" journalists' *Law Society Gazette* (Dublin 17 November 2018).

43 *Cornec* (n 34 above) [42].

44 Ibid [61].

45 Carolan (n 7 above) 185.

46 See generally ibid.

1 There is no legally recognised journalist privilege against non-disclosure of sources in Irish law (all authorities).
2 Irish law recognises a (democratically necessitated) presumptive public interest in confidential source protection under both article 40.6.1º of the Constitution (*Cornec*) and article 10 of the ECHR (*Cornec* and *Mahon v Keena*). An appropriate shorthand for this interest can be called 'journalist privilege'.
3 Both legal routes – article 10 ECHR, or article 40.6.1º – place significant weight on this public interest in source protection, but it is not absolute: an 'overriding' countervailing public interest can 'pierce' this 'privilege' (*Mahon v Keena* and *Cornec*).
4 The article 40.6.1º privilege is open to non-journalist 'publishers' fulfilling the democratic and educative function envisaged in the Constitution (*Cornec*).
5 The article 40.6.1º privilege applies to both the identity of the source and the information imparted to journalists/publishers by that source (*Cornec*). The question of whether identity or information is protected (or both) is determined by what the source expected to be kept confidential when they disclosed (*Cornec*).
6 The decision as to whether this privilege (either article 10 or, we can assume, article 40.6.1º) applies, or can be pierced, is a matter for the 'administration of justice'. The balancing of competing interests must therefore, in accordance with article 34.1 of the Constitution, be done by a court. Attempts by journalists to circumvent this judicial role can be punished (potentially by cost orders, and, we might assume, findings of contempt) (*Mahon v Keena*).

DEVELOPMENTS IN 2020: *FINE POINT FILMS* AND *CORCORAN*

This section deals with two significant judgments on journalistic privilege handed down within two months of each other in Northern Ireland and the Republic.

In a rare coincidence of timing, the Northern Irish and Irish High Courts in both *Fine Point Films* and *Corcoran* were each asked to examine if judges in lower courts had properly authorised search warrants for a journalist's home. In both cases, the focal issue was whether those authorising judges had properly considered the rights of journalists to resist disclosure under article 10 ECHR and, in the case of *Corcoran*, article 40.6.1º Bunreacht na hÉireann.

Alongside the novel coincidence and legal issues at play in both cases, it is useful to contrast legal developments across the Irish border when those developments are so closely connected in timing and facts.

In practical terms, it is also helpful for both lawyers and journalists to compare the approach in the neighbouring jurisdictions on the island. After all, many individuals from both professional camps work cross-border.

Most significantly, the applicant journalist in *Corcoran* also sought to use the judgment in *Fine Point Films* as supporting authority in their own challenge.[47] This strategy bore more fruit in the subsequent Court of Appeal judgment than it had in the High Court.

In Re Fine Point Films & Others

The Northern Ireland High Court judgment in *Fine Point Films* was delivered in July 2020. The background to this judicial review centred around the controversial documentary *No Stone Unturned*.[48] The film investigated allegations of British state collusion (through the policing and intelligence estates) in the 1994 'Loughinisland massacre', where six civilians were murdered by loyalist paramilitaries. After its release, the Police Ombudsman for Northern Ireland became aware that the documentary-makers had gained access to sensitive information from the Ombudsman's own examination of the original investigation into the killings by the Royal Ulster Constabulary. The film described this information as having come from an 'anonymous source' in 2011. The Ombudsman reported their suspicions to the Police Service of Northern Ireland (PSNI), who sought the assistance of Durham Constabulary in investigating how the documentary-makers acquired the information: 'whether by theft or other unauthorised disclosure'.[49]

The key legal dispute in the judicial review centred around the legality of the *ex parte* procedure adopted by the District Judge in deciding to grant a search warrant to the investigating officer from Durham Constabulary. The latter sought the warrant in order to recover the relevant documents and/or search for evidence of how, and from whom, they came into the possession of the film-makers. This included identifying who had leaked the material.[50] The particular warrant-granting power at issue is governed by the above-noted section 10 of the Contempt of Court Act 1981. The procedures governing this power under section 10 were contained in the Police and Criminal Evidence (PACE) (Northern Ireland) Order 1989: specifically articles 10, 11, 13 and 15, and schedule 1 (paragraphs 3, 9, 11). The absence of a comparable statutory regime in the Republic, and its in-built

47 *Corcoran* [2022] (n 5 above) [81].

48 Director Alex Gibney (Fine Point Films, 27 September 2017).

49 *Fine Point Films* (n 1 above) [13]. Counsel for the police also suggested – in the High Court's view without evidence – that a criminal breach of the Official Secrets Act 1989 was in play (at [35]).

50 Ibid [27].

protections for journalists, was, as we shall see, a key issue in the subsequent *Corcoran* judgments.

When seeking the search warrant, Durham Constabulary had argued before the District Judge that an *inter partes* hearing involving the journalist (as was required under schedule 1, paragraph 11(d) where 'journalistic material' was being sought) should be dispensed with in that instance. The police argued this was required given the highly sensitive nature of the documents, and the alleged risks to life their publicity posed. The investigating officers also argued that an *ex parte* hearing was necessary to secure potential evidence, as the journalists involved had previously refused to voluntarily disclose information to the police on the basis of an asserted journalist privilege. Their commitment to the privilege, the investigators believed, suggested these journalists might destroy the documents if they were on notice of the intention of the police to search for them.[51] Without referring to the article 10 ECHR rights of the journalists, the police (with whom the District Judge agreed) argued:

> The public interest was asserted to be the benefit to the investigation from the retrieval of the information which would help protect life, prevent and deter crime and restore and maintain public confidence within law enforcement.[52]

In overturning the decision of the District Judge to grant the warrant without an *inter partes* hearing, the Divisional High Court began its review by framing the granting of such search warrants as a 'draconian power' and 'a nuclear weapon':[53] particularly in the context of journalists' article 10 ECHR rights.[54] The court went on to find that the *ex parte* hearing was, given the significance of the power granted, procedurally unfair. Citing Hughes LJ in *In Re Stanford International Bank Ltd*,[55] the court found that, in order to be fair, such a hearing requires the applicant

> ... to put on a defence hat and ask, what, if he were representing the defendant or a third party with a relevant interest, he would be saying to the judge. The applicant must, of course, then proceed to tell the judge what those matters are. It is against that standard that we have reviewed the conduct of the application and hearing in this case.[56]

---

51 Ibid [37].

52 Ibid [39].

53 Ibid [22], quoting *R (Faisaltex Ltd) v Crown Court at Preston* [2009] 1 WLR 1687, [29] and *R (Mercury Tax Group) v Revenue and Customs Comrs* [2009] STC 743, [71].

54 *Fine Point Films* (n 1 above) [21], citing *Roeman and Schmidt v Luxembourg* (2003) ECtHR 51772/99.

55 [2010] 3 WLR 941.

56 *Fine Point Films* (n 1 above) [42].

By not raising the article 10 ECHR rights of the journalists before the District Judge – particularly the *Goodwin* right to source confidentiality, and its requirement of an 'overriding requirement in the public interest' to pierce it – the District Judge was unable to undertake the appropriate balancing of interests required under the Convention.[57] The failure to provide evidence to support the claim that the journalists might destroy the evidence was additionally fatal to the fairness of the procedures. Indeed, the court found the police's casting such crude aspersions on journalistic ethics could, in and of itself, 'completely undermine the important role that journalistic sources play in our democratic society'.[58]

Though unnecessary for their determination of the case, in his conclusion Morgan LJ for the court went beyond the narrow procedural point at issue and found that the available materials demonstrated 'no overriding requirement in the public interest which could have justified an interference with the protection of journalistic sources in this case'.[59] This barred further attempts by the police to secure a search warrant on the applicant's premises using the same evidential basis.

All in all, *Fine Point Films* is a vociferous, full-throated endorsement of the central value of journalist source protection in a democracy and a strict application of article 10 ECHR to the powers of police search and seizure in UK law: one welcomed by journalists.[60] The court's careful emphasis on the significance of such powers of search by the state, and the stand-alone importance and value of journalistic autonomy in a healthy democracy, were of particular note.

Before moving on to the *Corcoran* judgment, it is worth reflecting on potential source motivations behind the disclosure of sensitive documents relating to 'Troubles'-era sectarian murders. Source motivation has become, as we shall see, an emergent threshold for article 10 ECHR protection. The Police Ombudsman, PSNI and Durham Constabulary were clearly of the view that the disclosure was likely the result of a theft from the Ombudsman, or breach of some legal duty of secrecy. The High Court seemed sceptical this was the case, but was, regardless, unperturbed by possible criminality on the part of the confidential source. The historic context of the Loughinisland killings, and significant political sensitivities around collusion in Northern Ireland, might permit us to assume that the Northern Irish High Court

57 Ibid [43].

58 Ibid [47]. This suggests that the Northern Irish court was ignorant of, or apathetic towards, the background in *Mahon v Keena* south of the border, where journalists had indeed destroyed documents for fear that they might be compelled by a court to disclose them.

59 *Fine Point Films* (n 1 above) [47].

60 See Jon Slattery, 'Sources court victory for Irish Times' (31 July 2009).

viewed the confidential source as comparable to a classic public-sector whistleblower: criminal or not.

While, to be clear, the question of source motivation was not at issue in *Fine Point Films*: the nature and context of the disclosure could easily be argued as fitting the paradigm of a 'noble' confidential source, motivated by police/state accountability. As we shall see in *Corcoran* and the ECtHR jurisprudence below, this is a paradigm with which the courts may feel more comfortable.

Corcoran v An Garda Síochána

Two months after *Fine Point Films*, the Irish High Court in *Corcoran* was asked to review a similar complaint by a journalist challenging the legality of a police search warrant. The police search at issue had resulted in the seizure of the journalist's mobile telephone by An Garda Síochána (AGS – the Irish national police). The case emerged from a highly publicised (in part thanks to the applicant journalist's reporting on the scene) attack on a farmhouse in Falsk in rural Ireland. In December 2018, in an isolated hinterland of Ireland's midlands, a family were forcefully evicted from their farm by a security company acting on behalf of a bank seeking to repossess the property following a High Court order. The family resisted, and their resulting violent removal (a particularly sensitive image in rural Ireland)[61] was captured on social media, and widely publicised in the surrounding region. The evicted family and their supporters eventually left the property, and the security company secured and remained in the house to ensure it was not re-occupied. That night, a group of approximately forty masked people attacked the house, driving the security workers from it, and committing a number of serious assaults and arson in the process. The applicant journalist Emmett Corcoran was a local newspaper reporter and editor who was contacted by an unnamed source. That source arranged to meet Corcoran and then escorted him to the farm during, or towards, the conclusion of the organised attack. Corcoran was the first journalist present, and he took a number of photographs and video footage of the scene on his phone. His reporting, which included images of burning vehicles outside the farmhouse, was quickly picked up by national news organisations, becoming a major news story.

Soon after, Gardaí (officers in AGS) interviewed Corcoran under caution as part of their investigation. Corcoran willingly shared copies of the footage he had gathered. However, he refused a request by Gardaí to inspect his phone in order to attempt to identify who had contacted him, citing his obligations to his sources under 'journalist

61 Rory Carroll, 'Masked vigilantes attack guards at Irish farmhouse after eviction: Roscommon eviction clash evokes Irish land struggle of 1800s' *The Guardian* (London 17 December 2018).

privilege'.[62] A few months later, Gardaí applied to the District Court for a search warrant of Corcoran's home and business premises under section 10 of the Criminal Justice (Miscellaneous Provisions) Act 1997. During the search Gardaí retrieved Corcoran's password-protected phone. However, before they were able to access its data – and begin identifying potential witnesses or suspects to the attack – Corcoran issued judicial review proceedings in the High Court challenging the legality of the search warrant.

Before Simons J in the High Court, Corcoran's complaint focused on the difficulty journalists claiming 'privilege' had in vindicating their rights under article 10 ECHR and article 40.6.1°[63] due to a lacuna in the Irish warrant-authorising legislation. Unlike in Northern Ireland – where, as discussed above, the PACE (Northern Ireland) Order 1989 made provision for *inter partes* hearings, enabling an 'independent person'[64] to assess if an 'overwhelming public interest' justified piercing a journalist's article 10 ECHR right to source confidentiality – the 1997 Act in the Republic does not provide for *inter partes* hearings in the District Court. Even more significantly, both parties before the High Court agreed that the 1997 Act also fails to grant the District Court jurisdiction to undertake an article 10 ECHR balancing exercise in determining whether to authorise a search warrant.[65]

According to Corcoran, this apparent lacuna in the statutory scheme for court-authorised warrants left journalists in his position with the sole recourse to the High Court to judicially review the Gardaí and the District Court in order to vindicate their constitutional and Convention rights. While in Corcoran's case the Gardaí had not yet accessed the relevant material as it was password-protected – giving him the time to seek injunctive relief against the state – the applicant noted this may not always be the case. Requiring resolution through High Court litigation would, in many instances, effectively compromise the journalist's rights if Gardaí could immediately access the relevant protected material during a search, which Simons J acknowledged.[66]

---

62 *Corcoran* [2022] (n 5 above) [15].

63 Simons J found that the same balancing approach as required under the Convention is also required under the Constitution: *Corcoran* [2020] (n 2 above) [49] and [60].

64 See *Sanoma* (n 33).

65 *Corcoran* [2022] (n 5 above) [60]. This striking conclusion appears to be based on an interpretation of 'organ of State' in Ireland's European Convention on Human Rights Act 2003 (that transposed the Convention into Irish law), which excluded the District Court. It is not clear, however, why counsel for *Corcoran* appears to have conceded that the District Court was similarly constrained from considering constitutional rights.

66 Ibid [107].

Corcoran's strategy of highlighting this alleged lacuna also had the benefit of reminding the courts of unresolved difficulties in applying a key component of the Supreme Court judgment in *Mahon v Keena*: namely statutory provision for the 'who decides' question.[67] Given the punitive forcefulness of the Supreme Court's determination in *Mahon v Keena* that this decision is the exclusive constitutional remit of the courts, this line of attack from Corcoran appeared, on its face, promising. The Supreme Court had, after all, criticised and punished journalists for robbing the courts of its rightful authority to determine if the privilege applied. Yet, according to Corcoran, the statutory scheme also effectively disabled the courts from doing so in failing to give the District Court the jurisdiction to consider journalists' article 10 ECHR and article 40.6.1º rights. Here, judicial authority was deprived by statute, rather than, as in *Mahon*, the journalist. The applicant buttressed this line of attack by pointing to the recent *Fine Point Films* judgment:[68] where, as we have seen, the proper functioning of the UK's statute-mandated *inter partes* hearing was held by the Northern Irish High Court to be crucial to article 10 ECHR vindication.

However, despite the apparent strength of this line of argument, Corcoran's counsel seemed to struggle over what appropriate (or strategic) remedy to pursue. One logical remedial solution to the applicant's argument might have been a finding of unconstitutionality. In failing to provide the District Court with the necessary power to undertake an article 40.6.1º/article 10 ECHR balancing exercise, the 1997 Act breached both the journalist's article 40.6.1º constitutional rights and constrained the proper judicial function under article 34.1.

Following this line of attack, the relevant part of the 1997 Act's warrant-authorising power would be declared unconstitutional, and therefore void *ab initio*. It would also be similarly inconsistent with the state's obligations under the Convention, codified in the European Convention on Human Rights Act 2003. Though, unlike the UK's Human Rights Act 1998, remedies under the 2003 Act are confined to the comparatively (to the Irish Constitution) toothless 'declaration of incompatibility'.[69]

While seeking a declaration of unconstitutionality might now appear an obvious, and desirable, remedy (a point noted by Simons J),[70] counsel for Corcoran either did not advert to it when drafting initial submission or decided against seeking it. Given the central importance of section 10 to day-to-day policing in Ireland, the latter approach may have been strategically advisable.

---

67 See *Keena v Ireland* (n 32).
68 *Corcoran* [2022] (n 5 above) [81].
69 Section 5, European Convention on Human Rights Act 2003.
70 *Corcoran* [2022] (n 5 above) [7].

Counsel for Corcoran did, for a time at least, argue for relief in the form of judicial reconstruction of the 1997 Act's provisions enabling *inter partes* hearings: a procedural safeguard which would, it was claimed, bring the Republic's regime closer into line with the Convention-compliant scheme in Northern Ireland. This remedy would have necessitated a significant expansion of the constitutional 'double-construction rule'[71] which enables Irish courts to read legislation in a manner consistent with the Constitution. However, Simons J judged this proposal too radical.[72] In any event, this line of argument seems to have been somewhat half-hearted, as counsel for Corcoran conceded during proceedings before Simons J that the District Court would still not have the jurisdiction to undertake the appropriate balancing of interests exercise required by both the Constitution and the Convention.[73] What good would a judicially constructed *inter partes* procedure be if the District Court hearing it could not weigh the interests of one party?

Instead, the applicant settled on the argument that the Gardaí should – in light of previous recognition of 'journalist privilege' by Irish courts – have sought their warrant before the High Court.[74] The Irish High Court's expansive constitutional jurisdiction[75] would have allowed it to undertake the necessary balancing of interests, thereby enabling the proper vindication of Corcoran's journalistic interests under article 40.6.1º and article 10 ECHR. If the High Court felt the 'privilege' could be pierced, order 50, rule 5 of the Rules of the Superior Courts[76] could, according to the applicant, have enabled the High Court to provide the Gardaí with the legal authorisation to search and seize the phone. The respondents' failure to seek this authorisation before the High Court meant that, according to the applicant, the warrant was unlawful, and Corcoran's phone should be returned by the Gardaí. As with *Fine Point Films*, this legal strategy saw journalists seek to use the 'who decides' question to attack Executive infringement. Given the examples of journalist ambivalence about the legal position on 'who decides', this must be treated as simple strategy, rather than a commitment to the legal boundaries of the privilege.

This proposed remedy appears to have satisfied whatever strategic concerns may have troubled the applicant, enabling Corcoran to avoid

71 *East Donegal Co-Operative Livestock Mart Ltd v Attorney General* [1970] IR 317.

72 *Corcoran* [2022] (n 5 above) [79], [81] and [82].

73 Ibid [60], [80] and [84].

74 Ibid [85].

75 Art 34.3.1, Bunreacht na hÉireann.

76 Order 50, r 5 reads, 'The Court … may make any order for the detention, preservation, or inspection of any property or thing'.

asking the High Court to find as unconstitutional key procedures for authorising routine police action. However, Simons J noted this argument also posed its own 'difficult legal issues'.[77] The absence of the Attorney General as a notice party in the case was, given the high constitutional stakes, additionally concerning for the judge.[78]

AGS's response was to argue that, as the matter was now before the High Court, Simons J could undertake the necessary balancing of interests, making the larger legal argument moot.[79] Simons J suggested this solution was also unsatisfactory,[80] perhaps perceiving the suggestion as a brazen attempt to avoid serious legal questions.

Despite this apparent desire to engage with the tricky constitutional issues, Simons J was able – with the assistance of counsel for AGS[81] – to identify his own pathway away from potential legal controversy. Simons J's neat solution was to deny that Corcoran's confidential source was actually a source at all – at least as far as the Constitution or the Convention were concerned:

> The Applicants' case is predicated on an assumption that … they are entitled to rely on journalistic privilege to resist disclosing the content of the mobile telephone. The Applicants' criticisms of the procedures adopted by An Garda Síochána all flow from that assumption. For the reasons which follow, I have concluded that that assumption is not well founded, and that there is no right to rely on a claim of journalistic privilege in this case.[82]

Simons J's pathway away from constitutionally and practically tricky waters was inspired by some recent developments in the ECtHR in relation to article 10 protections for journalist privilege where the court attributes particular nefarious characteristics or intention to the confidential sources.

While the applicant in *Corcoran* was, perhaps understandably, operating under the assumption that crime reporters might legitimately have, as part of their rolodex of confidential sources, those actually engaged in crime: Simons J fastened onto a little-discussed[83] ECtHR decision from 2014 to help narrowly resolve the case.

---

77 *Corcoran* [2022] (n 5 above) [88].
78 Ibid [89].
79 Ibid [87].
80 Ibid [88].
81 Ibid [101].
82 Ibid [90].
83 With the exception of 'Journalistic sources: right to receive and impart information' (2014) 5 European Human Rights Law Review 537.

A source-motive test for confidential source protection?

Stichting Ostade Blade v Netherlands[84] involved an article 8 and 10 challenge by a magazine publisher in the Netherlands to a domestic court's authorisation of a police search and seizure of computers. The search was part of an investigation into a bombing campaign by environmental activists in the 1990s. The magazine in question, *Ravage*, had suggested it would publish a letter from the group claiming responsibility. This claim attracted the attention of the investigating police, who, under the supervision of an investigating judge, attempted to find the letter in warrant-authorised searches of the magazine's premises. The magazine sued for breach of their Convention rights under articles 8 and 10, but the domestic courts rejected their claims.[85]

In determining that the magazine's subsequent appeal to the ECtHR was inadmissible, that court set out a significant new dimension to the *Goodwin*-line[86] of article 10 jurisprudence on journalist source protection. In *Stichting Ostade Blade*, the Strasbourg court focused on the motivations (or, more accurately, the motivations the court speculated were operative) behind the eco-terrorist source's decision to contact the magazine. It found their ambition was solely to enhance publicity for their bombing campaign: a sinister appropriation of the magazine's power to communicate with its readership. The court used this determination of nefarious source motive to establish a new evaluative threshold for article 10's confidential source protection. The ECtHR in *Stichting Ostade Blade* concluded that article 10 did not apply in this case, as the public did not have an interest in knowing the information generated from a cynical exploitation of both the news media's communication power and the privileges article 10 provides them. Citing the idea of confidential 'sources in the traditional sense' from its earlier decision in *Nordisk Film & TV A/S v Denmark*,[87] the effect of *Stichting Ostade Blade* is to erect a distinction between confidential news sources worthy of article 10 protection and those deemed unworthy.

For the High Court in *Corcoran*, this refinement of the scope of article 10 allows a court to dispense with the need to balance the interest in criminal investigation against the overriding interest that

84 (2014) 59 EHRR SE9.

85 Interestingly, the Amsterdam Court of Appeal did find that while the attempt to identify the journalists' sources by searching their premises was not an art 10 breach, investigating potential links between the magazine and the environmental activist group (ie a 'fishing expedition' for evidence based solely on the magazine's reporting) was a breach of art 10. See *Stichting Ostade Blade* (n 84) [27].

86 *Goodwin* (n 3 above).

87 40485/02 (Decision 8 December 2005) section I.

normally attaches to journalism. On this view, *Stichting Ostade Blade* positions a source motive test *before* the article 10-mandated *Goodwin* balancing of interests takes place.

The democratic value of hearing from nefariously motivated sources is presumptively deemed weaker than the state's interest in criminal prosecution: a position of questionable merit, given the democratic value surely comes from the information released, not the motivation of the messenger. While it is beyond the scope of this article to examine the implications of this position in detail, it suffices to say that, so interpreted, *Stichting Ostade Blade*'s source motive test may have a significant impact on the capacity of crime-reporting journalism to legally resist disclosure of confidential source material.

While the High Court in *Corcoran* structured its analysis and reasoning in line with the article 10 jurisprudence, the case was expressly determined under article 40.6.1º.[88] This codification of ECtHR jurisprudence directly into the Constitution was, in and of itself, quite significant. However, given the textual distinctiveness of article 40.6.1º, it is unfortunate Simons J was seemingly unwilling to consider if article 40.6.1º offered journalists different, or perhaps more expansive protection – as Hogan J had suggested in *Cornec* – to that provided under article 10 ECHR. For example, it was open to the High Court to hold that *Stichting Ostade Blade* excessively and illegitimately narrows confidential source protection, and that article 40.6.1º excluded this recent Strasbourg development.

The High Court in *Corcoran* began the decisive part of its judgment by comfortably attributing a nefarious motive to the applicant's anonymous source. Simons J concluded that the source's desire was to publicise their criminal wrongdoing (as the ECtHR found in *Stichting Ostade Blade*), and, more seriously, to intimidate future security workers who might attempt to undertake evictions on behalf of financial institutions.[89] This 'not unreasonable inference' is critical to the High Court's finding that the quality of the journalist–source relationship here – where the source sought to exploit the mass communication power of the journalist to disrupt otherwise lawful conduct – did not

88 The High Court judgment in *Corcoran* [2020] (n 2 above) began by emphasising the supremacy of the Irish Constitution over the Convention in this legal dispute, but then proceeded to rely exclusively on authorities from the latter. The effect is to silently transpose the Strasbourg court's recent narrowing of art 10 jurisprudence into art 40.6.1º of the Irish Constitution. This privileging of the Constitution over the Convention is mandated under the Irish constitutional order: if a legal dispute can be decided under Irish constitutional rights and norms, that resolution takes precedence over any potential remedies under the Convention, or the 2003 Act that codified it in Irish law. See *Carmody v Minister for Justice, Equality and Law Reform* [2009] IESC 71.

89 *Corcoran* [2022] (n 5 above) [92] and [99].

attract article 40.6.1º protection. Indeed, not only was this kind of confidential source relationship unworthy of constitutional protection, such a source was also – *per Stichting Ostade Blade* – not *really* a confidential journalistic source at all.[90] This finding dispensed with the need for an enhanced weighting for a journalist's confidential source interest in any balancing exercise by the courts.

To reinforce his findings regarding the source's nefarious motive, Simons J also included a purported balancing of interests analysis, which both minimised the extent of the encroachment on Corcoran's article 40.6.1º rights, while emphasising the significance of the state's interest (*qua* 'public interest'):

> An Garda Síochána seek to conduct a very limited examination of the content of the mobile phone in support of their investigation of alleged offences of the most serious kind.[91]

It is perhaps trite to observe that describing the encroachment sought by An Garda Síochána as a 'very limited examination' neatly elides how the examination in question – whose exclusive purpose was to identify Corcoran's source – fundamentally compromises the applicant's asserted confidential source interest. In terms of the constitutional rights at play, the examination sought by the Gardaí was, in truth, of grave significance for the applicant, his professional work and the ethical framework which governed that work.

In its final substantive conclusion, the High Court again returned to the question of source motivation. Drawing on *Stichting Ostade Blade*, Simons J found that the source was not 'motivated by the desire to provide information which the public were entitled to know'.[92] Because of this absence of proper motivation – regardless, it seems, of whether the information provided *was* of the kind the public were entitled to know – the source relationship did not attract the protection of journalist privilege.

A potential wider scope of *Corcoran*: no protection for crime-reporting journalism?

Though the High Court's finding that Corcoran's source was not *really* a source seemed designed to avoid the kind of robust weighting for confidentiality required by *Goodwin*, the judgment still ended up engaging in a lot of balancing. While these aspects of Simons J's judgment may be confined to *obiter*, they do offer some interesting insights into in what circumstances Irish courts might apply journalist privilege where a criminal investigation is at issue.

90 Ibid [100].
91 Ibid [93].
92 Ibid [102].

> The disclosure of journalistic sources might well be disproportionate in the case of *minor offences*. This would be especially so where the alleged criminality relied upon is directly connected to the publication complained of. More specifically, an allegation that the provision of information to a journalist had involved the 'theft' of confidential information would not, generally, be sufficient to defeat a claim of journalistic privilege. Were it otherwise, it would be all too easy to suppress the publication of material by conjuring up an alleged criminal offence.
>
> The facts of the present case are, however, entirely different. Here, the criminal conduct alleged consists of the carrying out of '*arrestable offences*' as defined. These are said to arise out of a serious assault and the destruction of property. The criminal conduct is extraneous to, and separate from, the disclosure or publication. I am satisfied that the public interest in ensuring that all relevant evidence is available in the pending criminal proceedings overrides the claim for journalistic privilege in this case ... there is [also] a related public interest in the proper investigation of criminal offences.[93]

These paragraphs raise two interesting points.

The first is the court's consideration of the connection between the transfer of information and the alleged criminal conduct. In what will no doubt be seen by journalists as a welcome move, this paragraph appears to provide confidentiality protection for journalists, where their sources have committed a criminal offence merely by sharing information. We might be forgiven for speculating that the judge had in mind certain paradigmatic kinds of public-sector whistleblowing: a topic of major public controversy and debate in Ireland over the past decade.[94] Though the judge did not directly connect the *Fine Point Films* case at this point, Simons J could well have argued *Corcoran* could be distinguished from the recent judgment from Northern Ireland, which, as we have seen, appeared to be that paradigm example of such a public-sector whistleblower.

The heavy emphasis on source motive in *Corcoran* and the supporting ECtHR authorities may raise questions about what other motivations the courts might deem unworthy: perhaps fame or financial gain.[95] The judgments in both *Stichting Ostade Blade* and *Corcoran* could be viewed as more concerned with the imputed source goal of attracting fame, rather than the source's criminal conduct. The

---

93 Ibid [95]–[97] (emphasis added).

94 See eg discussions by Charleton J in *The Disclosures Tribunal: Third Interim Report* (November 2017).

95 For example, the offence under s 62 of the Garda Síochána Act 2005 requires that the person leaking the information knows it will have a harmful effect. The court's emphasis on motive here might suggest art 40.6.1 protections would not apply where the information was leaked in contravention of s 62.

difficulty with how these courts have privileged source motive over the substantive quality and value of the information disclosed is that it is not clear what the boundaries on legitimate source motivation are. Whistleblowers may disclose for a variety of reasons: but their value to healthy democratic governance is arguably not rooted in their motivations, but the accuracy and insightfulness of the information they bring to light.

Potentially of far greater significance is Simons J's distinction between Garda investigations into 'minor' and 'arrestable' offences. The former, it seems (perhaps depending on motive), attract the protection of journalist privilege under the Constitution. However, where Gardaí are investigating the more serious category of 'arrestable offences',[96] journalists should not, it seems, expect to successfully raise the privilege. This, it is suggested, is a step beyond the source motive threshold for article 10 protection set out in *Nordisk Film* and *Stichting Ostade Blade* and narrows journalist privilege under article 40.6.1º even further.[97] This aspect of *Corcoran* may create significant challenges and risks for crime reporters who use confidential sources to report on any serious criminal wrongdoing. Simons J's reasoning here does not suggest it is confined to circumstances where the source themselves are the perpetrators of the said arrestable offence. That the confidential source may know materially relevant information about the alleged 'arrestable offence' could, it seems, suffice for piercing journalist privilege in favour of a criminal investigation. This may, in effect, functionally exclude much crime and policing reporting from article 40.6.1º protection.

Simons J considered that the court in *Re O'Kelly* had gone 'too far' by concluding journalists had no special position in the Irish Constitution order.[98] However, his judgment in *Corcoran* privileges the 'public interest in the proper investigation of criminal offences',[99] with little consideration for any distinct value of crime reporting in Irish democracy. This may suggest judicial scepticism about any such value.

---

96 Defined under s 2 of the Criminal Law Act 1997 to include a broad array of offences under Irish criminal law.

97 Indeed, if this aspect of *Corcoran* holds sway, it may confine the constitutional protection Simons J was willing to afford to the noble whistleblowers, where the said source commits an 'arrestable offence' by transferring information: for example, s 9 of the Official Secrets Act 1963.

98 *Corcoran* [2022] (n 5 above) [94].

99 Ibid [97].

2022: COURT OF APPEAL JUDGMENT IN *CORCORAN*

Following a post-judgment re-hearing of *Corcoran* before the High Court, where it became clear to the parties and the court that the judge's original order for data disclosure from Corcoran's phone was not technically possible, the High Court issued a supplementary judgment on the scope of disclosure in the case.[100] The core of this ruling was to exclude from the data disclosed to AGS the 'contact details (such as names, telephone numbers, email addresses etc) saved on the mobile telephone'.[101] The apparent effect of this ruling was highly significant given the general tenor of the original judgment. If no names, phone numbers or email addresses could be disclosed to AGS, then the disclosure would not provide the investigating officers with the specific information they sought: namely some identifying information (in the form of a name, a number or an email address) from Corcoran's source in the Falsk incident. While the High Court seemed unwilling – in principle at least – to vindicate the journalist's specific claim of privilege in the case, that court appeared willing to so vindicate through the practical effect of its disclosure order.

Both sides brought appeals of these judgments: the applicant journalist focusing on the legal principle, and AGS on the practical effect of the High Court's order.

The Court of Appeal in *Corcoran*[102] – in a significant judgment that sets out a strong constitutional status for journalistic privilege in Ireland – took the opportunity to revisit some of the key assumptions made by counsel and the High Court in the original judgment. Most notably, the court here reopened the question of the jurisdiction of the District Court to consider journalistic privilege when deciding to grant a search warrant under section 10 of the Criminal Justice (Miscellaneous Provisions) Act 1997. As noted above, the High Court found – and the applicant's counsel apparently agreed – that the District Court did not possess the necessary jurisdiction under section 10 to undertake the balancing of interests required in assessing a claim of journalistic privilege. The Court of Appeal found this conclusion was made in error.[103] According to the court, this error was based on an incorrect belief by counsel and the High Court that the right of journalistic privilege always required an *inter partes* hearing. This is, according to the court, not always the case – sometimes *ex parte* hearings can, once managed appropriately, vindicate the rights of the journalists. This is a point made clear by the court in *Fine Point Films*.

100 *Corcoran v AGS* [2021] IEHC 11.
101 Ibid [4].
102 *Corcoran v AGS* [2022] IECA 98.
103 Ibid [128]–[129].

District Court jurisdiction to consider journalistic privilege

On the question of District Court jurisdiction, the Court of Appeal delved into the jurisprudence on the relationship between administrative and judicial functions/powers.[104] Here, the court concluded that while section 10's power was an 'administrative act' – not the 'administration of justice' – section 10's designated decision-maker (the District Court) was nevertheless 'acting judicially' in exercising that warrant-granting power.[105] In 'acting judicially' then, the District Court was *required* to consider the impact on fundamental rights that might flow from a decision to grant AGS a search warrant for a journalist's home and work premises. The court found the failure of the District Court to appropriately consider such rights in its decision to grant the warrant resulted in an unlawful interference with the applicant's rights under the Constitution and Convention.

The Court of Appeal's judgment then went on to significantly develop the procedural parameters governing warrant-granting in cases of journalists in Ireland. The court agreed with the High Court in *Corcoran* that section 10 did not permit *inter partes* hearings along the lines provided for under Northern Ireland's statutory scheme. However, Costello J's judgment for the Court of Appeal went on to set out the procedural requirements the District Court should have observed in the *ex parte* hearing. Approving, and drawing significantly upon the Northern Ireland High Court's judgment in *Fine Point Films*, the Court of Appeal found that part of AGS's 'duty of candour'[106] to the District Court in that case included disclosing that the journalist had asserted privilege over the material.[107] More importantly, AGS were also constitutionally obligated to advise the District Court of the significant rights and interests in play under the Constitution and Convention in granting a search warrant of a journalist's home and workplace. At a 'minimum', AGS should advise the District Court:

> ... the warrant may result in the seizure of material captured by journalistic privilege, if the judge is advised of his or her obligation to take account of this in issuing the warrant, and if a legally sufficient basis for overriding that privilege is identified and explained.[108]

Following this required minimum disclosure by AGS to the District Court, it was then for that court 'to determine whether [AGS] had convincingly established that there was an overriding requirement in the public interest which required that the journalistic privilege

104 Ibid [99]–[105].
105 Ibid [100], [108]–[110].
106 Ibid [106].
107 Ibid [133]–[135].
108 Ibid [100], [111]–[113].

should be interfered with and they should be compelled to reveal their sources'.[109]

Oireachtas on notice

Throughout the Court of Appeal's analysis of the facts and law, a number of warnings were directed towards the Oireachtas about the gaps and inadequacies in Ireland's section 10 warrant-granting regime. The Supreme Court judgment in *Corcoran* also echoed these warnings. Among the more subtle of these saw the court echo comments from the Northern Irish court in *Fine Point Films* about the preferability of disclosure orders to search warrants in terms of risk to constitutional and Convention rights.[110] That the Irish statutory scheme makes no provision for such disclosure orders in criminal investigations was the first of two major weaknesses highlighted.[111] As we shall see below, the judgment seems to suggest this state of affairs will make it hard for AGS to get source information that journalists have asserted privilege over, even where that source is, *per Stichting Ostade Blade*, not a real source.

The more serious warning shot to the legislature centred around the Court of Appeal's repeated reference to the 'appropriateness' of *ex parte* hearings in determining whether to grant warrants against journalists. The court made clear that – though not in this particular case – there would be instances where an *ex parte* hearing was inappropriate in the circumstances, and that the Constitution required an *inter partes* hearing to vindicate the journalist's rights:

> It may well be that there will be circumstances in which under Irish law it is not appropriate that this exercise be conducted on foot of an *ex parte* application and, to that extent, Irish law is deficient in failing to provide for a procedure of the kind considered in *Fine Point Films* whereby an *inter partes* hearing can be conducted while at the same time enabling protection of the information against destruction pending that hearing.[112]

The obvious potential outcome of such a case where an *ex parte* hearing was inappropriate would be a finding of unconstitutionality against the section 10 warrant-granting regime and a striking-down of the law: a result with potentially significant consequences for policing in Ireland. The court concluded by calling for the enactment of a scheme comparable to that north of the border:

> ... it would undoubtedly be preferable if the Oireachtas legislated in this complex area and established a clear constitutional and conventional

109 Ibid [138].
110 Ibid [119].
111 Ibid [148].
112 Ibid [124].

compliant procedure analogous to that in the Criminal Justice (Miscellaneous Provisions) Act, 2011 in respect of legal professional privilege or that which exists in Northern Ireland.[113]

Application of *Stichting Ostade Blade*

The Irish Court of Appeal also employed an interesting application of *Stichting Ostade Blade*. The court accepted, along the lines discussed by the High Court, the Strasbourg case as authority for some journalist sources not being 'real' sources for the purposes of article 10 ECHR and article 40.6.1º.[114] Remember, the High Court in *Corcoran* had used *Stichting Ostade Blade* to avoid the potentially messy constitutional issues that lurked beneath this case. If Corcoran's source was not a 'real source', then the journalist's rights were not engaged, and there was no need to worry about whether the District Court had properly considered such rights in granting the warrant. In contrast, the Court of Appeal found that, even if Corcoran's source was not a 'real source', the rights of the journalist were still engaged and should have been weighed by the District Court. The Supreme Court in *Corcoran* roundly endorsed this approach.

Part of the reason for this departure from the High Court's conclusion in *Corcoran* appears to be that, given the wide breadth of the search warrant granted to AGS, other legitimate sources could easily be identified to Gardaí during their search. Even where the weight to be attached to the journalist's interest in such a case was much less than for legitimate sources, a balance must be struck to ensure no legitimate sources are identified. The Court of Appeal also, unlike the High Court, rejected the claim by AGS that the appeal court could retrospectively legitimate the warrant by undertaking its own balancing exercise. Once the original warrant-issuing court had failed to consider the journalist's rights, the constitutional breach had occurred and could not be remedied.[115]

If the Court of Appeal was seriously concerned about the expansive breadth of search warrants, and their potential to breach constitutional rights, this again highlights the weakness of the Irish regime in failing to empower the courts to grant disclosure orders in such criminal investigations. The court's focus on the 'sledgehammer' nature and effect of search warrants also potentially neuters the kinds of negative ramifications of *Stichting Ostade Blade* for crime-reporting journalism in Ireland identified in this article: at least until the Oireachtas legislates for disclosure orders.

---

113 Ibid [147].
114 Ibid [97].
115 Ibid [97].

However, there is another possible interpretation of the Court of Appeal judgment on this point. The general tenor of Costello J's findings focused on the procedural failures in the District Court in failing to balance the rights. While the court accepted the Strasbourg authority, it spent no time reflecting on the actual public interest weighting that should have attached to Corcoran's confidential sources. Indeed, as we have seen, the court rejected the appropriateness of such an *ex post facto* weighing of interests by an appeals court. If the District Court in *Corcoran* had actually engaged in the balancing of interests, and concluded that the source was not a source 'in the traditional sense', then it is difficult to imagine what grounds for complaint the Court of Appeal would have. If this is the more accurate interpretation of the Court of Appeal's judgment in *Corcoran*, then the problems identified here for some forms of crime reporting likely persist. Indeed, the concurring judgments of Collins J and Hogan J in the *Corcoran* appeal to the Supreme Court, indicate this is the case.

Implications for the Garda Síochána (Powers) Bill

This key part of the judgment also, intentionally or not, contained another potential warning shot to the Irish legislature. In response to recommendations by the Commission on the Future of Policing in Ireland, the Department of Justice has proposed two significant pieces of policing legislation which are before the Oireachtas. One, the Garda Síochána (Powers) Bill 2021, includes a controversial[116] proposal for a criminal offence where a person refuses a request to divulge passwords for their digital devices.[117] While the Court of Appeal was at pains to emphasise that password protection should not be relied upon by the state to retrospectively legitimise an unlawful search and seizure, the facts of *Corcoran* make clear how important maintaining password protection can be to give practical effect to journalistic privilege. If, for example, Corcoran had felt compelled to disclose the password for his phone to Gardaí out of fear of criminal sanction – conceivably giving investigating Gardaí immediate access to his confidential source information – then it is unlikely that the subsequent reviewing courts could have effectively vindicated his rights. The Court of Appeal judgment in *Corcoran* suggests, at the very least, that this proposed offence requires very careful consideration to avoid improper encroachment on fundamental rights.

116 See Irish Council for Civil Liberties, 'Power to compel passwords must be removed from Police Powers Bill' (1 June 2022).

117 Head 16 – Powers under search warrant.

EMERGING PRINCIPLES GOVERNING JOURNALISTIC PRIVILEGE IN IRELAND (AND EUROPE)

The Court of Appeal in *Corcoran* took some time to set out in detail the relevant principles (28 in all) governing journalistic privilege in Ireland.[118] The initial six principles set out at the beginning of this article are all included in the Court of Appeal's restatement. However, a number of new additions and refinements have been made in the recent judgments. All have been described and analysed thus far in this article.

In general terms, *Corcoran* has continued the High Court in *Cornec*'s broad importation of the *Goodwin*-line of article 10 ECHR jurisprudence directly into article 40.6.1º Bunreacht na hÉireann. The wholesale importation of ECtHR jurisprudence on article 10 ECHR's confidential source protection into article 40.6.1º is, to say the least, notable. The Court of Appeal *Corcoran* judgment in particular suggests a newfound appetite in the Irish appellate court to enforce robust protections for confidential news sources: one which will surely be welcomed by Irish journalists. Like the Northern Irish High Court's judgment in *Fine Point Films*, the Court of Appeal's judgment in *Corcoran* is a full-throated endorsement of journalist privilege when facing the coercive and punitive arms of the state. The judgment is the first where the courts have actively constrained part of a serious criminal investigation to protect the rights of journalists to protect confidential sources. The recent Supreme Court judgment in *Corcoran* has, however, raised some significant doubts about the *Cornec* line of article 40.6.1º authority. Here, the Supreme Court was careful to emphasise that its ruling was narrowly focused on article 10 ECHR principles, and the 2003 ECHR Act. While Hogan J consciously and thoughtfully reaffirmed his position in *Cornec*, the *obiter* judgments of O'Donnell CJ and Collins J make clear that the constitutional status of 'journalistic privilege' in Ireland remains an open question. Both judgments preferred to avoid conclusive determinations on the question until the point had been fully argued at trial. Collins J's judgment was, however, tonally sceptical regarding the veracity of the *Cornec* line of authority.

This article has also suggested that there remains ongoing uncertainty about the potential scope of *Stichting Ostade Blade* in Irish and European law. This article argues that this authority potentially undermines confidential source protection for some kinds of crime-reporting journalism. One interpretation of *Corcoran* implies that warrants against journalists in Ireland will always face the enhanced protection afforded to confidential new sources under article 10 ECHR

118 *Corcoran* [2022] (n 102 above) [97].

and article 40.6.1º. On this view, journalistic privilege should not be pierced even when a police investigation into a serious crime is in play. Though the Northern Irish court in *Fine Point Films* did not consider the implications of *Stichting Ostade Blade*, that court's willingness to similarly nullify a core component of a police investigation into a serious crime suggests a similar approach will be adopted there.

However, this article also notes that this interpretation of the judgment in *Corcoran* may not give comfort to crime-reporting journalism. The 'source motive' test described here that emerges from *Stichting Ostade Blade* has now been affirmed as a part of the Irish Constitution's freedom of expression protections. The Court of Appeal accepted that very little weight might be attached to sources the court deems ignoble.[119] The primary concern in *Corcoran* was, after all, procedural.

For the police and warrant-authorising courts, these recent cases have demonstrated that appellate courts in Ireland are now willing to nullify police search and seizures where fundamental rights are ignored in the warrant-granting procedure. Police seeking such warrants must now carefully emphasise to the authorising court that the subject of the search is a journalist; that confidential source material may be gathered during that search; and that such confidential source material carries heightened constitutional protections. These judgments also suggest that warrant-authorising courts will have to demonstrate clearly they have given the necessary consideration and weighting to journalistic privilege in deciding to grant a search warrant. Both Irish and Northern Irish courts have found that *ex parte* hearings can, once they observe the necessary consideration of fundamental rights, vindicate those rights. However, there are certain – as yet undefined – circumstances where an *inter partes* hearing is necessary under both the Convention and Constitution.

CONCLUSION

This article has sought to disentangle a number of complex, interconnected features of Irish and European legal protection for confidential journalistic sources. Trends on both sides of the Irish border over the past two years suggest that Irish courts are now taking journalistic privilege seriously and are willing to arrest and nullify coercive police and state encroachment on journalistic autonomy. These judgments are notable for providing robust endorsements of journalistic freedom and an appreciation of the practical challenges in

119 The judgments of both Collins J and Hogan J in the Supreme Court also unproblematically endorsed *Stichting Ostade Blade*.

realising those freedoms. *Fine Point Films* and the Court of Appeal in *Corcoran* saw Irish courts set firm boundaries on the scope of criminal investigations and issue new robust guidance to warrant-authorising courts on the necessity, value and importance of journalists' rights in our democracies. The Court of Appeal in *Corcoran* also issued some stark warnings to the Oireachtas regarding current deficiencies in Irish law, with implications for proposed enhancements of police power and criminal law.

There remain, however, some unresolved aspects of these cases. In particular, the High Court judgment in *Corcoran* suggests serious potential impacts of *Stichting Ostade Blade* on crime-reporting journalism. Though the Court of Appeal successfully avoided some of the inevitable implications of that Strasbourg authority, the quick willingness of both courts to dismiss the potential value of so-called 'non-traditional' sources in our democracies is worth interrogating more closely.

Northern Ireland
Legal Quarterly
Summer Vol. 74.2 (2023) 269–302
Articles DOI: 10.53386/nilq.v74i2.1096

The scope and rationale(s) of the change of position defence

Duncan Sheehan

University of Leeds
Correspondence email: d.k.sheehan@leeds.ac.uk

ABSTRACT

The article examines an innovative suggested rationale for change of position – namely that the claimant has 'outcome responsibility' for the defendant's change of position. It concludes that the justification fails. Although it purports to justify a single baseline against which to judge if the defendant's position has changed, it – at best – only justifies a subset of the cases in which change of position is normatively attractive; it does not justify the defence in (say) cases of innocent wrongdoing. As such it requires us to accept that there are several different species of defence. An easier route to justifying the availability of the defence in all these different cases is 'irreversible detriment', although that explanation still has to justify why the defendant should not be worse off.

Keywords: change of position; unjust enrichment; irreversible detriment; disenrichment; outcome responsibility; restitution for wrongs.

INTRODUCTION

The defence of change of position was introduced into English law by *Lipkin Gorman v Karpnale*.[1] Lord Goff suggested that it would apply in cases where the defendant had so changed his position that it would be unconscionable to make the defendant repay or repay in full.[2] The question for those researching the defence has therefore been: when is repayment unconscionable? The typical case is where the defendant has detrimentally relied on the receipt in making extraordinary expenditure that is now irreversible and has done so in good faith. In *Lipkin* itself, Cass, a partner at the claimant firm, stole money and gambled it away at the defendant's casino. Change of position succeeded because the club had changed its position in paying out Cass's winnings in reliance on the receipt of the initial stakes.[3]

1 [1991] 2 AC 548.

2 Ibid 577–583; other cases have also rested relief on inequitability. See *Garland v Consumer Gas* 2004 SCC 25, [2004] 1 SCR 629, [64]; *Dextra Bank v Bank of Jamaica* [2001] UKPC 45; [2001] 1 All ER (Comm) 195, 204.

3 [1991] 2 AC 548, 582–583.

The justification for change of position is contested. Often restitution lawyers speak about its rationale as being one of disenrichment.[4] That is, the defendant was enriched, but no longer is. Others have defended the defence on the basis that the defendant requires security of receipt. Put differently, the defendant should be able to rely on an assumption that he is entitled to keep – and use – money or assets received, unless he is disqualified from that assumption, most obviously because he knows differently.[5] James Edelman – now Edelman J of the High Court of Australia – has persuasively argued that security of receipt, however, is the result of having the defence not the reason for it,[6] and we do not pursue this idea here. Another issue arises as to whether change of position is properly a defence or a denial.[7] A denial is a claim that one or more of the prerequisites for the cause of action have not been satisfied. Disenrichment as a rationale may imply that change of position is a denial; if the defendant is not enriched, the requirements of the cause of action are not met. A defence by contrast acknowledges the completeness of the action but alleges a different reason to protect the defendant; it is exculpatory. A defeat provides a third possibility. A defeat[8] does not allege that the cause of action is not made out, nor does it really allege that the defendant has an excuse for not making restitution. Rather it undercuts the rationale for the action without precisely denying any of its elements.

The article's central aim is to assess the rationale for change of position initially put forward by Ajay Ratan[9] compared to other rationales put forward, principally disenrichment and irreversible detriment. Ratan argues that the reason that the defendant can avail herself of the defence is that the claimant is in some sense responsible for the change in her position. We explore in detail how this argument works later, but for now it suffices to note that it cannot justify the whole range of cases where the defence might arguably be said to apply. It may justify the defence in cases of personal unjust enrichment claims where the defendant has relied in some way on receipt. That is the classic instance of the defence. It could potentially justify the

4 Peter Birks, *Unjust Enrichment* 2nd edn (Clarendon Press 2005) 209–210; Harry Liu, 'Changing the shape of change of position' (2004) 15 King's Law Journal 301

5 Charles Mitchell, Paul Mitchell and Stephen Watterson (eds), *Goff and Jones: The Law of Unjust Enrichment* 9th edn (Sweet & Maxwell 2016) para 27.41

6 *Australian Financial Services & Leasing Pty Ltd v Hills Industries Pty Ltd* [2014] HCA 14, (2014) 253 CLR 580, [92] (Crennan J et al).

7 Andrew Dyson et al, 'Introduction' in Andrew Dyson, James Goudkamp and Frederick Wilmot-Smith (eds), *Defences in Unjust Enrichment* (Hart 2016) 1.

8 Dennis Klimchuk, 'What kind of defence is change of position?' in Dyson et al (eds) (n 7 above) 69, 85.

9 Ajay Ratan, 'The unity of pre-receipt and post-receipt detriment' in Dyson et al (eds) (n 7 above) 87.

defence in some, but not all, proprietary cases, where the defendant is the initial recipient of the property and has relied on receipt in some way. The justification has real difficulty in explaining the defence's application to cases where a third party steals the enrichment from the defendant. There are also reasons to believe that change of position should apply in some wrongs cases, allowing the tortfeasor to reduce his liability in restitutionary damages. However, we will see that Ratan's justification cannot apply in the same way as in the unjust enrichment context. Change of position in the wrongs' context must on Ratan's view be a different defence. If remote recipients can avail themselves of the defence in a tracing claim made against them, this must also be a separate defence on Ratan's view. The article therefore suggests that Ratan's view implies there may be at least two and quite possibly more than two different change of position defences.

This article is structured as follows. In the first part, we outline some of the differing rationales of change of position which have been offered, beginning with disenrichment. This was rejected by Australian case law and by cogent arguments put by Elise Bant. However, the irreversible detriment view she propounds, which relies on the defendant's complaint that she would be worse off if made to repay the claimant – and this is no less true of the disenrichment thesis – requires us to decide against which baseline we measure if the defendant is worse off – the 'no worse off than what?' question. The irreversible detriment view is intuitively attractive, supported by authority, and Bant has ably worked through the implications of the justification. Importantly for our purposes the irreversible detriment view does not preclude the availability of change of position in wrongs or proprietary claims, and we explore this in detail in the second part of the article. However, Bant fails to fully explain why the defendant should not be made worse off than the *status quo ante*. That failure leads us to Ratan's paper where he seeks an explanation to justify his suggested baseline that the defendant be no worse off than had the claimant not made (for example) the mistake. In the second part we explore how widely the change of position defence should apply and whether the scope of its application can be explained via Ratan's thesis. We suggest that in the end the thesis Ratan propounds, while interesting, begs more questions than it answers. Proffered as a means of explaining a choice of baseline, it could fill the gap in Bant's thesis even if such was not Ratan's explicit intention. It, however, fails to do so and should be rejected. It does not even plausibly explain the reach of the defence across all unjust enrichment claims. The reach of a 'no worse off thesis' is broader than outcome responsibility can justify. If Bant's explanation of a unitary defence of change of position is to succeed, another justification of why the defendant should not be made worse off is needed.

JUSTIFICATIONS POSITED FOR THE DEFENCE OF CHANGE OF POSITION

The defence has developed significantly over the years, and this has rendered it difficult to identify a clear single unique rationale. Originally it was analysed through the lens of a mistake claim; the suggestion was that when relief for mistake was relatively restricted, succeeding only in cases of liability mistakes, the defence was not needed but, if all causal mistakes allow relief, the defence was required to cut back restitution.[10]

Its scope has since narrowed and broadened. It has narrowed in that it may not apply to duress claims because of the defendant's actions and fault in inducing his own unjust enrichment; although there seems little in the way of authority on this point[11] to the extent to which bad faith (say) is an indicator of economic duress,[12] the defence should be excluded. The defence, as we see, is said not to apply if the defendant changes his position in bad faith, knowing he is not entitled to the enrichment; it is hard to see many cases where a duressor would not know he was not entitled. Possibly change of position does not apply to failure of consideration claims, or at least not all such. In *Haugusund Kommune v Depfa ACS Bank*,[13] for example, the Court of Appeal had to decide whether the local authorities, having entered a void swaps agreement with the banks, were able to rely on change of position in answer to a claim for restitution. They were not able to do so. Aikens LJ drew a distinction between two cases. The first is where the defendant receives money believing it is hers to keep and the second is where the defendant receives money, knowing that she will have to repay it at some point.[14] In such a case the defendant cannot rely on a change of position defence to justify a refusal to repay because the enrichment was accepted on the basis that it would have to be repaid/paid for. On the facts Aikens LJ held that the authorities did take the money on the basis that it was repayable. The agreements were always void,

10 *Barclays Bank v WJ Simms & Sons Ltd* [1980] QB 677, 695-696 (Goff J); Peter Birks, 'Change of position and surviving enrichment' in William Swadling (ed), *The Limits of Restitutionary Claims* (UK National Committee of Comparative Law 1997) 36, 40–41.

11 Duncan Sheehan, 'Defendant-sided unjust factors' (2016) 36 Legal Studies 415; Andrew Burrows, *The Law of Restitution* 3rd edn (Oxford University Press 2011) 544–545; there are a number of undue influence cases where the defendant influencer was innocent and able to rely on the defence. See *Cheese v Thomas* [1994] 1 WLR 129.

12 *D&C Builders v Rees* [1966] QB 617; *Pakistan International Airlines Co v Time Travel (UK) Ltd* [2021] UKSC 40, [56]–[59] (Lord Hodge et al), [102] (Lord Burrows).

13 [2010] EWCA Civ 579, [2012] QB 549; *Goss v Chilcott* [1996] AC 788.

14 [2010] EWCA Civ 579, [2012] QB 549 [123].

but that 'cannot change the basis on which the kommunes received the money'.[15] Burrows – now Lord Burrows of the UK Supreme Court – makes a convincing argument that, while this holds in loan cases, it does not hold in other scenarios. His example is a builder who receives money in advance, pays for material and for a holiday. The builder does not take on any obligation to repay money, but to do the work. If there is a failure of consideration, he ought to be able to rely on the holiday expenditure to found the defence.[16] At the same time the defence has broadened; for example, some cases indicate that it may apply to innocent wrongdoing[17] where the defendant is unaware of the wrong, often a trespass or conversion.

It has also fractured internally in case law or commentary in two ways. First, Birks contrasted disenriching cases with alleged non-disenriching cases where the defendant is not financially worse off but is still deemed to have changed his position, although it seems extremely difficult to identify clear case law examples of the latter. The second line of fracture contrasts the reliance cases and non-reliance cases. In typical reliance cases the defendant relies on the receipt of the enrichment in making a decision to dissipate it; on the narrow view of the defence this is a necessary condition of its applicability, and there was formerly some dispute as to whether such reliance could take place in anticipation of receipt. It is clear now that it can.[18] Detrimental reliance does not require a link to be proven between specific receipts and specific items of expenditure;[19] general reliance on increased assets will suffice.[20] The defendant is disqualified from relying on the defence if she had no legitimate expectation of being able to rely on the receipt of the asset. This might be because the defendant knew, or was

15 Ibid [124]

16 Burrows (n 11 above) 544–545; Andrew Burrows, *A Restatement of the English Law of Unjust Enrichment* (Oxford University Press 2012) §23(2)(iii), 117–122; Stevens claims this is still incompatible with the basis of the claim. Robert Stevens, 'The unjust enrichment disaster' (2018) 134 Law Quarterly Review 574, 587. But see *Benzline Auto Pte Ltd v Supercars Lorinser Pte Ltd* [2018] SGCA 2, [2018] 1 SLR 239; Charles Mitchell and Paul Mitchell, 'Recurring Issues in failure of basis' [2020] Lloyd's Maritime and Commercial Law Quarterly 498, 509.

17 *Cavenagh Investment Pte v Kaushik Rajiv* [2013] SGHC 45 (trespass to land); *Kuwait Airways v Iraqi Airways* [2002] 2 AC 883, [79] (Lord Nicholls) (conversion).

18 *Lipkin Gorman v Karpnale* [1991] 2 AC 548; *Dextra Bank v Bank of Jamaica* [2001] UKPC 45; [2001] 1 All ER (Comm) 195.

19 Paying off debts will not suffice therefore; *Scottish Equitable BS v Derby* [2001] 3 All ER 818; *Credit Suisse v Attar* [2004] EWHC 374 (Comm), [98]; *Goff and Jones* (n 5 above) para 27.11.

20 *Philip Davis v Collins* [2000] 3 All ER 818; *Skyring v Greenwood* (1825) 4 B&C 281, 107 ER 1064; *Holt v Markham* [1923] 1 KB 504.

wilfully blind to the fact that,[21] the claimant had made a mistake or was in bad faith.[22] It may be sufficient that the defendant should have made further inquiries as to her entitlement,[23] although it is clear that English law does not ask about the relative fault of the parties and bar the defence if the defendant is more at fault.[24] Bad faith cancels out the initial assumption of reliance[25] and demonstrates that the defendant's actions were not caused by the receipt of the money. In the second set of cases the change in the defendant's position is independently caused by a third party or act of God. The wide view of the defence, accepted in England,[26] accommodates this; the narrow view of the defence does not. We look in turn at three possible explanations for change of position: disenrichment, irreversible detriment and outcome responsibility. We deal with them in this order because, arguably, it is defects in the previous explanations that spawned the later ones.

Disenrichment and unjust disenrichment

Will a rationale of disenrichment work? There is a pleasing logic to it. If the cause of action responds to the fact that the defendant was enriched initially, the defence should respond to the way that enrichment falls away subsequently. In other words the rationale is related to the initial reason for having a claim. Restitution is available against the defendant to the extent – and only to the extent – that it removes sufficient enrichment to return the defendant to the *status quo ante*. Removing more renders the defendant worse off than previously. This justification has attracted some level of judicial support. In *Test Claimants in the FII Litigation v HMRC*,[27] Henderson J said that the defence was 'essentially concerned with disenrichment'. However, change of position is not, in Mitchell and Goudkamp's language, a

21 *Harrison v Madejski* [2014] EWCA Civ 361, [61]; *Port of Brisbane Authority v ANZ Securities* [2003] 2 Qd R 661, 674–675.

22 *Lipkin Gorman v Karpnale* [1991] 2 AC 548 (HL) 577.

23 *Jones v Churcher* [2009] EWHC 722, [46]; *Niru Battery Manufacturing v Milestone Trading* [2004] QB 985; *State Bank of NSW v SBC* (1995) 39 NSWLR 350; *Heperu Pty Ltd v Belle* [2009] NSWCA 84, (2009) 76 NSWLR 195; *Citigroup v National Australia Bank* [2012] NSWCA 381, (2012) 82 NSWLR 391; *Cavenagh Investments Pte Ltd v Kaushik* [2013] SGHC 45, [2013] 2 SLR 543, [71] (Chan Seng Onn J). For comment, see eg Robert Chambers, 'Change of position on the faith of receipt' [1996] Restitution Law Review 103, 107–108 and *Goff and Jones* (n 5 above) paras 27.41–27.44

24 *Dextra Bank v Bank of Jamaica* [2001] 1 All ER (Comm) 195.

25 Jessica Palmer, 'Chasing a Will-o'-the-Wisp: making sense of bad faith and wrongdoers in change of position' [2005] Restitution Law Review 53.

26 Andrew Burrows, 'Change of position: the view from England' (2003) 36 Loyola of Los Angeles Law Review 803; *Scottish Equitable BS v Derby* [2001] 3 All ER 818.

27 [2014] EWHC 4302, [2015] STC 1471, [354].

denial.[28] It is not simply a statement that an aspect of the cause of action is not proven. It is not therefore right to treat change of position as part of the general enrichment enquiry.

There are three reasons. First, we remove cases of bad faith disenrichment from the defence;[29] not all causal disenrichments therefore count.[30] The immediate problem therefore with disenrichment as a rationale is that it does not explain why bad faith disenrichments for instance are excluded, and this links to a point below that disenrichment provides no normative basis by itself for an exculpatory defence. We could base a denial on disenrichment, but this would involve our measuring the defendant's enrichment at the time of trial; that we do not do this is the second reason to reject change of position as part of the general enrichment inquiry. A rationale based on disenrichment is incomplete and over-inclusive.

The third reason is that the disenrichment view may also be under-inclusive, excluding non-disenriching changes of position. A position must be taken by those who put weight on disenrichment whether to bar such changes or to explain the defence's availability in such cases differently. Birks took the former view. He argued that there were two different and distinct defences,[31] although saying that non-disenriching changes of position would be rare.[32] Discussion has revolved around the case of *Commerzbank v Price-Jones*.[33] That was a case of a foregone financial benefit. Price-Jones had deliberately chosen not to seek higher-paying employment elsewhere; the court decided that this was too speculative and too evidentially uncertain to succeed, and the decision does not make for a good test case for this reason. In *Palmer v Blue Circle Southern Cement Ltd*,[34] again there was a foregone financial benefit. The defendant chose not to apply for social security benefits to which he was otherwise entitled on the basis of the receipt of the money. Bell J decided that that could be an example of a change of position, following an English estoppel decision to come

28 Charles Mitchell and James Goudkamp, 'Defences and denials in the law of unjust enrichment' in William Swadling (ed), *Restatement, the Third, of Restitution and Unjust Enrichment: Critical Essays* (Hart 2013) 133, 156–157; *Goff and Jones* (n 5 above) para 27.06.

29 James Edelman, 'Change of position: a defence of unjust disenrichment' (2012) 96 Boston University Law Review 1009, 1020–1021.

30 Burrows (n 11 above) 526.

31 Birks (n 4 above) 258–261

32 Peter Birks, 'Change of position: two central questions' (2004) 120 Law Quarterly Review 373, 378.

33 [2003] EWCA Civ 1664; see also *Kinlan v Crimmin* [2006] EWHC 779, [2007] 2 BCC 102.

34 [1999] NSWSC 697.

to this conclusion,[35] and this decision seems not to be a good test case either. One reason is that the change of position in *Palmer* is reducible to money. In fact most alleged non-disenriching changes of position seem ultimately reducible to money. Another reason is the reliance on an estoppel case when, as we see, estoppel may be based on a separate rationale. Other non-disenriching changes of position mooted include going to university; the defendant in *Gertsch v Atsas*[36] gave up work to do so, foregoing income (although possibly raising future earnings power). Another is having a child. Bant has suggested that this latter example cannot be included in disenrichment without stretching the idea altogether out of recognition.[37]

Australian High Court jurisprudence has also accepted that some changes cannot be included in disenrichment and has put forward an alternative rationale – irreversible detriment – discussed in detail in the next subsection. In *Australian Financial Services Pty Ltd v Hills Industries*,[38] AFSL was induced by a fraudster to make payments to Hills for non-existent equipment and to enter into leaseback arrangements regarding the equipment with companies owned by the fraudster. Hills treated the payment, as requested by the fraudster, as discharging certain debts owing to them from other companies, themselves owned by the fraudster, and refrained from taking action against them. The High Court emphatically rejected disenrichment as a rationale for change of position. This was in large measure precisely because some relevant changes of position will be difficult or impossible to value.[39] The plurality also criticised 'disenrichment' as being overly mathematical when the law should ask who should bear the loss and why.[40] French CJ argued that disenrichment was, at best, a circumstance defining a class of case in which recovery could be held inequitable and founded the defence very firmly on 'a general rubric of inequitable recovery'[41] as set out in *Lipkin Gorman v Karpnale* and subsequently in *Dextra Bank v Bank of Jamaica*.[42] More broadly the court did not therefore adopt an unjust enrichment analysis of restitution, and so the apparent symmetry of enrichment versus disenrichment alluded to earlier simply did not arise. Instead

35 *Avon CC v Hewlett* [1983] 1 WLR 605.
36 [1999] NSWSC 898.
37 Elise Bant, *The Change of Position Defence* (Hart 2009) 143; Elise Bant, 'Change of position: outstanding issues' in Dyson et al (eds) (n 7 above) 133, 142–143.
38 [2014] HCA 14, (2014) 253 CLR 580.
39 Ibid [23] (French CJ).
40 Ibid [78], [84].
41 Ibid [23] (French CJ); [79]–[80] (Hayne J et al); [144]–[145] (Gageler J).
42 [2002] 1 All ER (Comm) 193 (PC).

the court[43] founded recovery on the equitable roots of the action in *Moses v Macferlan*,[44] something to which both French CJ and the plurality referred in discussing the foundations of the defence. Gageler J described change of position as being the second stage of an analysis based on 'notions of conscience'.[45] Accepting that some hard-to-value detriments should count, disenrichment seems under-inclusive. In general, of course, in order to calculate the reduction in liability the change of position must be reduced to money,[46] and yet Bant's point above about children – and therefore the more general point in *AFSL* – holds. It is difficult to see a child purely in terms of the cost of milk, nappies and baby food, and the defence could be seen as absolute in exceptional cases if the qualifying change of position cannot be valued. Giving some support for this in *Kinlan v Crimmin* Deputy Judge Sales commented:

> *Even if he may still have in his hands the monies paid to him* ... Mr Crimmin changed his position in a fundamental respect ... Had he realised that the agreement was invalid and the payments made under it were made by mistake, Mr Crimmin would obviously have wished to consider how his continuing interest in the company should be protected, either by his resuming his rights to protect himself as a quasi-partner in the business or by seeking the reformulation of the agreement so as to ensure that it and the payments to him were valid. These opportunities which were denied him cannot be restored to him.[47] (emphasis my own)

The important point here is that no financial detriment is necessary. The defence is available on this view even if the defendant still has the money (the enrichment) paid. Indeed, it is clear from *RBC Dominion Securities v Dawson*[48] that the defendant was enriched by the value of the furniture bought with the mistaken payment. Change of position still applied despite the extant enrichment. The mere fact she still benefited from the money did not defeat the change in her position.

We have seen now that disenrichment is under-inclusive. The first reason for rejecting disenrichment *per se* was that it was over-inclusive because it could not explain why bad faith defendants were excluded from the use of the defence. We can put this objection in a different way; disenrichment *per se* cannot provide a normative reason for an exculpatory defence. On one view, change of position concerns

43 [2014] HCA 14, (2014) 234 CLR 580, [65]–[76] (Hayne J et al); [105]–[126] (Gageler J).

44 (1760) 2 Burr 1005, 97 ER 676.

45 [2014] HCA 14, (2014) 234 CLR 580, [143].

46 *Goff and Jones* (n 5 above) para 27.31.

47 [2006] EWHC 779 (Ch), [2007] BCC 106, 121-122 (Sales DJHC).

48 (1994) 111 DLR (4th) 230.

defendant autonomy, a suggestion made by Lord Reed in *Benedetti v Sawiris*.[49] This is sometimes said to be linked to the enrichment enquiry. To explain the argument, Burrows endorses a link, originally made by Peter Birks, between change of position and subjective devaluation, explicitly developed in the context of the enrichment criterion,[50] and it was in the context of a discussion of subjective devaluation that Lord Reed made his suggestion as to the relevance of autonomy to change of position. Subjective devaluation, as applied to the cause of action, is intended to protect the defendant's freedom of choice by allowing him to argue that he would not have obtained it at a given market price – ie it is worth less to the defendant than the market price and the defendant is not enriched to the same extent. I might say that, although the market price for having my house painted was £1000, I would never have agreed to have it painted magnolia, and so the house painting is only worth £100 to me. In the context of change of position, I might fairly say that I would only have sought the particular service I bought after I received the enrichment. Without that enrichment I would never have spent the money and requiring me to retransfer the money with no credit is tantamount to forcing me to pay for an unwanted service, which subjective devaluation says I should not be forced to do.[51]

This actually suggests that the rationale for the defence is that the defendant's decision to change her position was vitiated by a mistaken belief or reliance on receipt – the 'unjust disenrichment view'.[52] In other words, the defendant spent this money which she would not otherwise have done in error. This is the standard case of change of position, but if the defendant transferred the money in error authority suggests she should recover against her transferee, the third party. In tax cases, for instance, Bant argues the recoverability of money by the defendant from the state bars the defendant's change of position defence.[53] Knox J, for example, said in *Hillsdown Holdings v Pensions Ombudsman* that the defendant's consequent liability to tax was not a change of position except 'to the extent Hillsdown is

49 [2013] UKSC 50, [2014] AC 938, [118], referring to *Sempra Metals Ltd v IRC* [2007] UKHL 34, [119].

50 Burrows (n 11 above) 527.

51 Peter Birks, *An Introduction to the Law of Restitution* revised edn (Clarendon Press 1989) 413.

52 Bant, *Change of Position Defence* (n 37 above) 144; see Edelman (n 29 above) for a developed view of this idea of 'unjust disenrichment'.

53 Bant, 'Change of position' (n 37 above) 144–145.

unable to recover the tax',[54] although Hillsdown's ability to do so was not litigated and Her Majesty's Revenue and Customs was not party to the action.

Disenrichment *per se* does, however, enable us to see a common link between the reliance cases and theft cases. In both the defendant is disenriched, providing a common link. This separate common link is vital since reliance is irrelevant to the latter set of cases. If the theft cases are included, causation might provide a necessary connection,[55] and English law does appear to be shifting to a view based on causation.[56] If the defendant's disenrichment were caused by the receipt, the defence will bite. One way of demonstrating such causation – but not the only one – is to point to the defendant's reliance on the receipt.[57] Causation is not normative, however; it is factual. To conclude, disenrichment does not fit the cases and cannot explain why some disenrichments do not count or why non-disenriching changes do count, assuming that they do. Secondly, by itself disenrichment does not explain why the *status quo ante*, as opposed to some other baseline, is appropriate.

Irreversible detriment

The Australian cases formulate a rationale of irreversible detriment. While this approach also asks whether the defendant is made worse off or not, the change of focus allows us to include non-disenriching changes of position, or cases where the defendant is seemingly still enriched.[58] As we saw in the previous section, this was one reason why the irreversibility criterion was authoritatively confirmed in *Australian Financial Services Pty Ltd v Hills Industries*, referencing the work of Elise Bant.[59] Another advantage of the approach over disenrichment is that it emphasises that detriment as assessed at the time of the claim rather than the time the change of position occurred. This is because the irreversibility criterion tells us that the defendant is not in a position to recover the money that he has paid away. Little is, of course, utterly

54 *Hillsdown Holdings plc v Pensions Ombudsman* [1997] 1 All ER 862, 904 (Knox J); *K&S Corp Pty Ltd v Sportingbet Australia Pty Ltd* [2003] SASC 96, (2003) 86 SASR 313; *Hinckley & Bosworth BC v Shaw* [1999] 1 LGLR 385.

55 Bant, 'Change of position' (n 37 above) 143–145.

56 *Philip Collins v Davis* [2000] 3 All ER 808; *Wards Solicitors v Hendawi* [2018] EWHC 1907, [32]–[33]; *Goff and Jones* (n 5 above) para 27.08, but see *Credit Suisse v Attar* [2004] EWHC 374 (Comm), [98].

57 Bant, *Change of Position Defence* (n 37 above) 153.

58 Bant provides a number of reasons why the irreversible detriment is preferable at ibid 134–138.

59 [2014] HCA 14, (2014) 234 CLR 580, [23]; see also *Alpha Wealth Financial Services Pty Ltd v Frankland River Olive Company Ltd* [2008] WASCA 119.

irreversible, but it appears purchases of, or improvements to, land may be taken as irreversible.[60] If litigation is needed for the defendant to recover money paid away, we might deem it irreversible.[61] Against that we can put the *dictum* of Knox J in *Hillsdown* that we saw earlier. It will be remembered that the fact the defendant had a right to recover the money he transferred in error from the Revenue rendered change of position unavailable. In practice it will be difficult to get the Revenue to repay in the absence of litigation, and so at best these two lines of authority sit uncomfortably.[62] Nonetheless, in *AFSL* itself the debts owed to Hills by the fraudster's companies were in effect unenforceable, and the change of Hill's position in giving up and discharging those debts was irreversible 'as a practical matter of business'.[63]

Potentially, the Australian position is narrower than the 'irreversible detriment' rationale implies. Australian courts have placed a great deal of emphasis on reliance.[64] The importance Australian courts place on reliance stems in part from a consideration of the relationship between change of position and estoppel. The analogy appears in several places in *AFSL*. The plurality,[65] for example, reference the decision in *Grundt v Great Boulder Gold Mines Pty Ltd*[66] on detriment in estoppel to bolster their case that what matters in change of position is detriment not disenrichment. Too much should not be made of this, however. Estoppel has a rather different focus and the High Court in *Australian Financial Services Pty Ltd v Hills Industries* seems, with respect, to misunderstand this. Gageler J saw change of position as being merely estoppel minus the representation[67] and appears to suggest on this basis first that change of position operates absolutely, unless that

60 *Saunders & Co v Hague* [2004] 2 NZLR 475; see, for discussion, eg Charles Mitchell, 'Change of position: the developing law' [2005] Lloyd's Maritime and Commercial Law Quarterly 168.

61 *K&S Corporation v Sportingbet (Australia) Pty Ltd* [2003] SASC 98, (2003) 86 SASR 313.

62 Bant, 'Change of position' (n 37 above) 145 refers to there being 'simple procedures' to recover from the taxing authorities and the transfer being reversible for that reason.

63 *Australian Financial Services Pty Ltd v Hills Industries* [2014] HCA 14, (2014) 253 CLR 580, [95].

64 *Optus Networks Pty Ltd v Gilsan (Intl) Ltd* [2006] NSWCA 171; *Ethnic Earth Pty Ltd v Quoin Technology Pty Ltd* [2006] SASC 7, (2006) 94 SASR 103; *David Securities v Commonwealth Bank of Australia* (1992) 175 CLR 353 (HCA), *Citigroup Pty Ltd v NAB* [2012] NSWCA 381, (2012) 82 NSWLR 391.

65 [2014] HCA 14, (2014) 253 CLR 580, [85]; [23] (French CJ).

66 (1937) 59 CLR 641.

67 [2014] HCA 14, (2014) 253 CLR 580, [155]–[158].

would be disproportionate.[68] This is a novel idea,[69] certainly if the absolute nature of change of position is to be the norm rather than an exceptional response to the 'unvaluability' of the defendant's change of position. Gageler J's point makes sense if change of position is indeed just estoppel minus the representation; however, as Hudson has pointed out, the idea behind estoppel is that a claimant who made a representation to the defendant that X was true is held to that where that fact is taken as the shared relevant state of things. The claimant is held to that because the defendant has relied to their detriment on the representation, and the claimant needs to be responsible for that. To allow one party to depart from the adopted state of affairs infringes the other's autonomy,[70] and so we hold the representing party to their representation.[71] The plurality in *AFSL* also comment that estoppel provides a level of protection to the defendant's expectations which change of position does not.[72] The tight connection with estoppel is also inconsistent with the view, accepted by Gageler J, that the payee can rely on information from sources other than the claimant payor.[73] While the full relationship between estoppel and change of position is beyond our scope, the point is that the justification for an all-or-nothing estoppel and a change of position defence are not the same. Estoppel is concerned with the defendant's autonomy; change of position with not rendering the defendant worse off.

The link with reliance and estoppel led Gageler J to exclude independent changes of position from the defence.[74] The question was not explored by all the justices. French CJ deliberately eschewed any analysis of the question.[75] There was reliance by the defendant and the question did not need to be decided. With respect, however, excluding such changes of position does seem to fly in the face of the stated rationale for the defence in *AFSL*. An independent change of position, such as the destruction or the theft of the asset, will be a detriment to the defendant. Requiring restitution of the value of a thing destroyed or stolen without fault renders the defendant worse off than had there been no transfer of the asset in the first place and, in practical terms,

68 Ibid [158]; see Bant, *Change of Position Defence* (n 37 above) 161; Bant, 'Change of position' (n 37 above) 160–162.

69 Bant, 'Change of position' (n 37 above) 161.

70 Jessica Hudson, 'The price of coherence in estoppel' (2017) 39 Sydney Law Review 1, 11–12

71 *Goff and Jones* (n 5 above) para 30.16; Bant, *Change of Position Defence* (n 37 above) 163.

72 [2014] HCA 14 (2014) 253 CLR 580, [86] (Crennan J et al).

73 Ibid [157] (Gageler J); *Citigroup v National Australia Bank* [2012] NSWCA 381, (2012) 82 NSWLR 392, [5] Bathurst CJ (et al).

74 Ibid [142].

75 Ibid [25].

that detriment is irreversible, and there is Australian authority – albeit first instance – for this.[76]

While this may be a legitimate criticism of *AFSL*, irreversible detriment has important advantages over disenrichment. It can – and by cutting the implicit link with the enrichment inquiry was designed to – accommodate non-disenriching and difficult to value changes of position.[77] It reflects a 'no worse off' rationale for the defence. While that rationale is also present in the disenrichment thesis, by cutting the link with the enrichment inquiry it leaves the door open to the defence's application in appropriate cases of restitution for wrongs and, indeed, also to proprietary claims, where appropriate. We can see the 'no worse off' rationale in the plurality's references to disadvantage resulting to the defendant if restitution were ordered[78] and more clearly in Gageler J's comment:

> The second condition [for the application of the defence] is that, by reason of having so acted or refrained from acting, the defendant would be placed in a worse position if ordered to make restitution of the payment than if the defendant had not received the payment at all.[79]

Change of position prevents the defendant from being in a worse – or entirely different – position after making restitution to the *status quo ante*. It is not the purpose of restitutionary remedies to do so.[80] The missing piece, as with the disenrichment thesis, is why this – as opposed to some other baseline – is appropriate.

Outcome responsibility

What is outcome responsibility?

In the introduction we suggested that an alternative justification – and one which was explicitly intended as clarifying the choice of baseline – might be that of Ajay Ratan. He identifies one potentially attractive way to proceed in justifying change of position, pointing to a link between the defence and foundational questions of the justification of liability, particularly in terms of ensuring the defendant is 'no worse off'.[81] His question therefore is 'no worse off than what?' It is the question of which baseline we use to assess whether and by how much the

76 In *Gertsch v Atsas* [1999] NSWSC 898 the theft of a luxury car was accepted as a relevant change of position. See also *Corporate Management Services (Australia) Pty Ltd v Abi-Arraj* [2000] NSWSC 361; Bant, *Change of Position Defence* (n 37 above) 136.

77 Bant, 'Change of position' (n 37 above) 147.

78 [2014] HCA 14, (2014) 253 CLR 580, [85].

79 Ibid [157].

80 Bant, *Change of Position Defence* (n 37 above) 171.

81 Ratan (n 9 above) 109

defendant has changed her position. There are a number of different possible baselines available. One we might call the 'no receipt' baseline. Here we assume that the change of position defence should not render the defendant worse off than if she had not received the enrichment.[82] Gageler J in *AFSL*, for example, quite clearly asks in the quotation just above whether the defendant would be worse off as against a baseline counterfactual that the defendant had not received the benefit: the 'no receipt' baseline. Yet this sits uneasily with the availability of the defence in cases of anticipatory reliance.[83] In anticipatory reliance cases the defendant had not received the enrichment when she relied. One solution is to tack on an additional baseline that the defendant be no worse off than had she not relied in anticipation of the receipt. Ratan's aim is to justify a single competitor baseline that operates in both anticipatory reliance and reliance *ex post* cases. He suggests that the defendant be no worse off than had the claimant not had his decision-making impaired.[84] He calls this the 'no defect' baseline[85] and argues that it does everything the 'no receipt' baseline does and more and is therefore preferable. We can agree with this, however, without necessarily accepting the outcome responsibility thesis he propounds.

One possible link between the two questions of the rationale for the claimant's ability to seek restitution in the first place and the rationale for the defence is to ask which baseline is the most compelling in justifying imposing liability. In other words in justifying awarding restitution at all we need to ask against which baseline is the defendant better off and remove just enough so that the defendant is no worse off; restitution is justified in the absence of defendant wrongdoing because the defendant will be 'no worse off'. This is another 'no worse off than what?' question. The same baseline can then be used in change of position. Ratan points to Grantham and Rickett's argument that corrective justice should be relevant to defences as an example of such a linkage between the rationales for liability and the defence. Grantham and Rickett explicitly adopt the corrective justice view of Ernest Weinrib,[86] although without requiring that it be the sole driver of all private law liability. Weinrib's view links Aristotelian corrective justice and Kantian right. Injustice occurs when the prior equality of the parties is disrupted – there is a breach of Kantian right. Both parties critically must be implicated. This is where the requirement of

82 Burrows (n 11 above) 528–531.
83 Ratan (n 9 above) 88.
84 Ibid 91.
85 Ibid 91–92.
86 Ross Grantham and Charles Rickett, 'A normative account of defences to restitutionary liability' [2008] Cambridge Law Journal 92, 98.

bilaterality comes from,[87] as well as the related requirement that the reasons for the claim must apply equally to claimant and defendant. Applying this to defences, Grantham and Rickett conclude that the reasons the defendant provides for excluding liability must also apply equally to the claimant and defendant.[88] Grantham and Rickett correctly sound a note of caution,[89] accepting that Weinrib never himself discussed defences in his treatment of unjust enrichment.

Ultimately, Ratan does not pursue a line of attack that links the fundamental rationale for requiring restitution with that of the defence of change of position. He thinks a justification of unjust enrichment along the 'no worse off' lines might not allow us to choose between baselines. This requires some explanation. Ratan takes the position that unjust enrichment theory has yet to provide a compelling case in favour of the 'no worse off' thesis. Wilmot-Smith indeed describes the whole argument as question-begging because there is no way of justifying eg a *status quo ante* baseline without answering the question why the defendant should not keep the transfer.[90]

Ratan takes a slightly different tack, appealing to outcome responsibility.[91] Outcome responsibility is in part constitutive of our identity. In short, if we are not responsible for our acts and their consequences on others, while there may be bodies and minds, there are no real people doing real things.[92] In wrongs cases, the wrongdoer is blameworthy in some respect and therefore the wrongful losses need to be repaired by the wrongdoer. Put differently the things we do are our responsibility not merely things that just happen. Robert Kane expresses it well,[93] saying that agents who express or exercise free will are authors of and characters in their own story. By virtue of self-forming judgements of the will, the agent is an arbiter of his own life, taking responsibility for making it what it is. Outcomes that we cause are ours in a way that those we do not so cause are not ours.[94] We therefore use outcome responsibility to morally attribute outcomes to agents. It is this that outcome responsibility adds to references to causation or disenrichment. Causation is factual. It provides no moral reason for the claimant's responsibility for the defendant's acts and

87 Ibid 100–102.
88 Ibid 104.
89 Ibid 105.
90 Frederick Wilmot-Smith, 'Should the payee pay?' (2017) 37 Oxford Journal of Legal Studies 844, 849–851.
91 Ratan (n 9 above) 104.
92 Tony Honoré, 'Responsibility and luck: the moral basis for strict liability' (1988) 104 Law Quarterly Review 530.
93 Robert Kane, 'Responsibility, luck and chance: reflections on free will and indeterminism' (1999) 96 Journal of Philosophy 217, 240.
94 Ratan (n 9 above) 104.

therefore no moral reason why the claimant should care. Outcome responsibility does provide a morally significant reason why the claimant should care what the defendant has done or suffered.

This requirement of moral attribution of blame may lie behind the relative fault requirements found in §142(3) Restatement of Restitution, and in §65 Restatement, the Third, of Restitution and Unjust Enrichment, comment f) of which says that a recipient whose negligence exceeds that of the claimant in the transaction by which the recipient was unjustly enriched cannot use the defence. There are signs of relative fault in Commonwealth case law as well. Relative fault can be used to attribute the loss,[95] and there are two ways to do this. First the defendant may be barred completely from availing herself of the defence should she be more at fault. This is the route of the Restatement. *Waitaki Intl Processing (NZ) Ltd v National Bank of NZ*[96] goes a different route. If the payer is thought to be 70 per cent at fault and payee 30 per cent at fault, only 70 per cent of the change of position can be available to the payee – ie the payee must absorb 30 per cent of his loss.[97] In *Waitaki* itself, the reduction was 10 per cent. The payor bank had continually insisted the payment was correct (which explains their 90 per cent responsibility); however, there were questions as to whether the account into which the payee put the money to keep it safe prior to repayment and the security for that was adequate. Henry J held it was not adequate and upheld the trial judge's allocation of 10 per cent responsibility to the payee.[98] This question of relative fault was to be judged in the 'round'. *Dextra Bank v BoJ* rejected both approaches to relative fault and subsequently Chisholm J accepted that as binding on him in New Zealand in *Saunders & Co v Hague*.[99]

Despite this, we might think that outcome responsibility lends itself well to a relative fault approach and *vice versa* – the relative fault approach could be justified as the court apportioning outcome responsibility. Ratan, however, does not take this route. He maintains that relative fault is the wrong way to think about things. On his view the effect of the mistake is to provide a justification for making

95 Scott Struan, 'Mistaken payments and the change of position defence: rare cases and elegance' (2012) 12 Otago Law Review 645, 653; See Henry Cohen, 'Change of position in quasi-contracts' (1932) 45 Harvard Law Review 1333, 1356–1358 to the effect that the meaning of predominant fault in this context is unclear.

96 [1999] 2 NZLR 211; *Thomas v Houston, Corbett & Co* [1969] NZLR 151; Ross Grantham and Charles Rickett, 'Change of position and balancing the equities' [1999] Restitution Law Review 157.

97 John McCamus, 'Rethinking section 142 of the Restatement: fault, bad faith and change of position' (2008) 65 Washington and Lee Law Review 889, 911–912.

98 [1999] 2 NZLR 211 (CA), 221–222; 229–231 (Thomas J).

99 [2004] 2 NZLR 475, 493.

restitution of the payment. If, however, the claimant's actions, albeit caused by his mistake, have led the defendant to rely on the faith of the receipt and pay money away, the claimant has outcome responsibility for that payment away if he can be said to have put the defendant in the position of believing himself entitled (irrespective of fault).[100] If the claimant cannot be said to have done that, the detriment cannot properly be brought into account.[101] There is no need to rely on relative fault. Outcome responsibility simply works on the basis that the claimant cannot expect the defendant to return the money or other asset, but at the same time not give credit to the defendant for actions (sufficiently) causally connected to his mistake. He must take the rough with the smooth.[102] The claimant cannot justify distinguishing between the consequences of the defect in her decision-making on the basis that one furthers her own interests (getting the money back) and the other does not.[103]

Moral luck plays an important role in this. Moral luck occurs when an agent is treated as an object of moral assessment despite a significant aspect of the moral judgment depending on factors beyond her control.[104] We are concerned here with resultant luck,[105] which occurs when our actions and projects turn out differently because of matters beyond our control. By paying over the money the claimant puts herself at risk of moral luck. This lies at the heart of the idea of respecting the claimant as a person. Her acts have consequences, and she has to live with them – not just some of them. Otherwise, she is not a real person doing real things. On this view, change of position is not merely a case of respecting the other's autonomy in the same way as you would expect yours to be respected (although arguably it may be that as well) because the claimant bears responsibility for the change of position. One rejoinder might be that the defendant is responsible for his actions not the claimant. We explore this later in the second subsection.

The defence responds to (in)action by the defendant. If the asset transferred is stolen, should the defence apply? It seems so. Birks put the point in this way. If change of position were not available, the receipt of anything would be a cause of dread, leading to the adoption of

100 See *Wards Solicitors v Hendawi* [2018] EWHC 1907, [36].

101 Ratan (n 9 above) 105.

102 Ibid 105.

103 Ibid 113; see also Tony Honoré, *Responsibility and Fault* (Hart 1999) 9, 1.34–135.

104 Dana K Nelkin, 'Moral luck', *Stanford Encyclopedia of Philosophy* (2019).

105 Thomas Nagel, 'Moral luck' in D Stateman (ed), *Moral Luck* (State University New York Press 1993) 57, 60.

extreme measures to protect assets.[106] A deliberate choice not to adopt these measures should see the defendant protected from liability just as a deliberate choice to pay away would entail his protection. Moral luck precisely brings into account things that are out of the claimant's control and so what the choice is that the defendant makes on account of the claimant's mistake does not matter. The claimant's outcome responsibility provides the relevant moral link justifying the presence of the defence. What, however, if the defendant makes no choice at all?[107] On one view even pure inaction can be covered. By putting the asset in the hands of the defendant, and therefore in a position to be stolen, the claimant bears some outcome responsibility. An obvious rejoinder is that the thief bears responsibility not the claimant, and we return to this in the next subsection, but it is worth saying a little here too. We may think pragmatically that the difficulty of distinguishing conscious negative reliance from pure inaction is too great and this might prove a persuasive reason allowing the defence in theft cases and other independent changes of position.

When are we outcome responsible?

We must therefore explore this idea of sufficient causal connection because, in the absence of a relative fault approach, it seems the only way to control for outcomes we are not responsible for as a claimant and which therefore cannot be brought into account by the defendant in change of position and to decide what outcomes the claimant is responsible for.[108] The normal test of causation in unjust enrichment cases is 'but-for', and there is authority that this holds true in change of position too. In other words, it is a necessary condition for the defence's application that but for the receipt or anticipated receipt the defendant would not have paid away the money,[109] despite some confusion caused by *dicta* by Mummery LJ in *Commerzbank* to the effect that the test was whether there was a relevant connection.[110] The connection on

106 Birks (n 4 above) 211; Liu (n 4 above) 304; *Scottish Equitable BS v Derby* [2001] 3 All ER 818.

107 This distinction between negative reliance and pure inaction is raised by Oliver Black, 'Varieties of legal reliance' (2017) 28 King's Law Journal 363, 377.

108 Ratan (n 9 above) 105 has the example of A setting fire to B's car and coincidentally also mistakenly paying B £100. This is slightly different in that the fire is not causally connected at all to the payment and so cannot be brought into account in availing B of a change of position defence justifying not returning the money. Here we are talking about how causally connected phenomena might still not be brought into account.

109 *RBC Dominion Securities v Dawson* (1994) 111 DLR (4th) 230; *Skandinaviska Enskilda Banken AB (Publ), Singapore Branch v Asia Pacific Breweries (Singapore) Pte Ltd* [2011] 3 SLR 540.

110 [2003] EWCA Civ 1663, [43] (Mummery LJ).

the facts was in fact a but-for link. Bant has suggested, however, that the appropriate test in unjust enrichment claims (or those related to impairments in the claimant's intention or decision) is not but for, but whether x was a factor in the claimant's decision.[111] This avoids the issue of over-determination[112] where there are several independently sufficient reasons/causes for an impugned decision. Bant takes the view that this 'a reason' test applies equally to the human decision-making process in change of position,[113] although counterfactual causation remains relevant to the non-reliance cases. In other words, the defendant must merely prove that the receipt of the money (or, on Ratan's view, the defect in the claimant's decision-making) was a factor in the decision to spend the money. Whichever view one takes – that the test is 'but-for' or 'a factor' – this still remains incomplete. We are not outcome-responsible for everything that happened or was decided because of (causally) our actions.[114] Responsibility is limited to outcomes properly attributable to the conduct. The idea of moral luck raised above does not preclude this; it does not require everything out of the claimant's control to be brought into account.

Let us start with cases where we are both outcome responsible.[115] Imagine A goes to a posh restaurant and spends B's mistakenly paid money. A is outcome responsible. A chose to go to the restaurant; A is responsible for that choice. However, it was a contributing factor to A's decision that B had given him the money. Consequently, B too could be outcome responsible. In every jurisdiction change of position seems uncontroversially available here. Yet A's outcome responsibility for his own actions seems to militate against this conclusion. A way out might be causal contribution. To what extent has B's action contributed to the loss? Honoré discusses the question of causal contribution[116] in *Responsibility and Fault*. His focus is tort, and in *Responsibility and Fault* Honoré applies causal contribution to contributory negligence, but by applying it in the change of position context we can test the workability of Ratan's hypothesis. Importantly, contributory negligence in tort requires an apportionment; the damages received by the claimant can be reduced to reflect her contribution to the loss.[117] That contributory approach has been rejected in change of position both by authority and by Ratan himself. Consistency with this

111 Elise Bant, 'Causation and scope of liability in unjust enrichment' [2009] Restitution Law Review 60.
112 Ibid 67.
113 Ibid 75–76.
114 Honoré (n 103 above) 77.
115 Ibid 89.
116 Ibid 91.
117 Law Reform (Contributory Negligence) Act 1945, s 1.

requires a more absolutist approach; questions of causal contribution become a gateway to the defence, and once through the gateway the defence is available to the full extent of the defendant's changed position. Hart and Honoré discuss causal contribution in the context of 'degrees of causation'[118] where we say that an outcome was caused partly by A but mostly by B.[119] The assessment of causal contribution may be rough-and-ready; Hart and Honoré refer to it as 'vague and commonsensical'.[120] On this view, if the party most responsible for the outcome is the claimant, the defence is available. If the party most responsible is the defendant, it is not.

In deciding this question of causal contribution and which party should bear the loss associated with the change of position, we could incorporate a normative judgment as to which causes count more than others. One way to do this might be by reference to whether the defendant has acted in reasonable reliance. If he has not reasonably relied, the defence ought not to be available.[121] On the outcome responsibility view, if the defendant has acted unreasonably in his reliance or in bad faith, the causal contribution of the claimant's (perhaps obvious) error is too low. A flipside example where the claimant does have sufficient responsibility may be this. There have been suggestions in Australian cases, albeit that Gageler J for example pulled back from this in *AFSL*,[122] to the effect that the information on which the defendant relies in making his decision to pay away should derive from the payer. This could be a central (but not the only) case where the claimant's responsibility is greater than the defendant's and where the defence should be available.

It will be necessary to decide when the claimant's outcome responsibility has effectively disappeared – ie where she is responsible for some of the defendant's changed position but not all of it. We may be looking for the unjust enrichment equivalent of a *novus actus interveniens*, or some type of remoteness rule.[123] The function of this rule would be to say when a change in the defendant's position is so remote from the original enrichment that the claimant cannot be

118 H L A Hart and Tony Honoré, *Causation in the Law* 2nd edn (Oxford University Press 1985) 233.

119 *Clay v Crump* [1963] 3 All ER 687 is a tort case where the court sets out three contributory factors in order of importance, and see Law Reform (Married Women and Tortfeasors) Act 1935, s 6(2).

120 Hart and Honoré (n 118 above) 233.

121 Bant (n 111 above) 77–78.

122 [2014] HCA 14, (2014) 234 CLR 580, [157], citing *Citigroup Pty Ltd v National Bank of Australia* [2012] NSWCA 381, (2012) 82 NSWLR 392, rejecting the view in *Swiss Bank Corporation v State Bank of NSW* (1995) 39 NSWLR 350.

123 On which see Richard Nolan, 'Change of position' in Peter Birks (ed), *Laundering and Tracing* (Clarendon Press 1995) 135, 149–151.

sensibly said to be responsible for the change. In tort law, a *novus actus interveniens* is an act of a third party breaking the chain of causation. Clerk and Lindsell observe that no precise test exists and refer to four issues: what was the impact of the intervening conduct, was the conduct deliberate or unreasonable, foreseeable and independent of the defendant?[124] In tort, remoteness rules attempt to provide for the reasonable foreseeability of the loss caused to the claimant so as to protect the defendant from being responsible for things which although counterfactually connected should not be brought into account.

There is relatively little general guidance in tort law, since the extent of the defendant's liability should reflect the policy behind the specific tort.[125] In some wrongs a no remoteness rule, or one allowing expansive recovery, might be appropriate, say in cases of deceit or fraud, where the defendant cannot be allowed to say the claimant's loss is too remote to be compensable.[126] In most tort cases the loss must be reasonably foreseeable. The unjust enrichment claimant is not even a wrongdoer, so, accepting *per arguendo* the tort analogy, we might argue that he should certainly not be taken to be responsible for anything more. Bant has, however, argued that remoteness rules are unnecessary and any causally related change should be taken into account. Indeed, this is one of her reasons why irreversible detriment is preferable to disenrichment, some of whose proponents have suggested a remoteness principle.[127] Some might also find the analogy with tort unconvincing. After all, part of the train of events in many unjust enrichment scenarios is precisely that nobody intended any of it.

However, even if remoteness rules are thought inapplicable, it will still be necessary to make a normative choice as to what operative causes count more than others in deciding whether the claimant or defendant should bear the loss. The very idea of claimant outcome responsibility suggests that the idea of *novus actus interveniens* is appropriate if the *novus actus* plausibly cuts the chain of responsibility. Where, for example, the money paid by the claimant is stolen by a third party from the defendant, we might conclude that it is the thief who is responsible, or maybe the defendant who fails to take precautions against the theft. If the claimant is not responsible (for whichever reason), it cannot be brought into account as a relevant

124 Michael A Jones, Anthony M Dugdale and Mark Simpson (eds), *Clerk and Lindsell on Torts* 23rd edn (Sweet & Maxwell 2020) para 2.114; *Chubb Fire Ltd v Vicar of Spalding* [2010] EWCA Civ 981.

125 *Clerk and Lindsell* (n 124 above) para 2.150.

126 *Doyle v Olby* [1969] 2 QB 158.

127 Bant, *Change of Position Defence* (n 37 above) 136, discussing eg Nolan (n 123 above).

change of position, but this seems very unfair on the defendant, particularly because it is precisely the need to avoid incentivising unnecessary precautions to protect assets that we saw Birks point to as a reason why we need the defence in this context.

EXTENDING CHANGE OF POSITION: DIFFERENT SPECIES OF DEFENCE?

There are a number of difficulties therefore with Ratan's proposed rationale. It will require us to construct an apparatus to assess causal contribution, but without that assessment becoming overly uncertain and 'hand-wavy' in the way that assessments of contributory or relative fault threatened to be overly uncertain. We have also seen that there is a real difficulty in including theft cases within Ratan's outcome responsibility justification for change of position and therefore within the scope of the defence. We might include it pragmatically, but it will seem a stretch to many to say that the claimant bears outcome responsibility, as a result of a sufficient causal connection, for the theft of the transferred asset. The analogy with tort which may help with 'causal contribution' in other ways seems to militate against it. Even, however, *per arguendo* accepting that theft cases can be justified via outcome responsibility, it is impossible to include change of position in wrongs within Ratan's justification, and there are difficulties in the application of his thesis to proprietary claims also. This section examines the justification for extending the defence to first wrongs and then proprietary claims and shows that, while an irreversible detriment view allows for the defence to apply in these cases, Ratan's outcome responsibility justification does not do so.

Extending the defence to innocent wrongdoing

This section explores the question whether the defence of change of position applies to wrongs – specifically to the cases usually dealt with under the heading of restitution for wrongs. These cases lie outside of unjust enrichment because there is a breach of duty by the defendant. There are some cases where disgorgement for a wrong is available, but no change of position applies, as for example where dishonest conduct is present,[128] which would bar the defendant from change of position as being in bad faith. There are wrongs of strict liability where use-damages are available against an innocent defendant where matters appear more open in principle. It is rarely if ever suggested that change of position should apply more widely. The application of the defence to wrongdoing appears to depend on tortious use-damages

128 Bant, *Change of Position* (n 37 above) 169.

cases being restitutionary in the sense of being gain-based rather than compensatory and loss-based.[129] If these damages are loss-based, change of position ought not to apply. The defence does not – in any of its guises – apply to claims to recover losses caused by the defendant. It is beyond the scope of the article to prove that these examples of use-damages are restitutionary, but they are commonly, although not universally, seen as such in the academy.[130] We start by examining the authority for the availability of the defence in this narrow context and then seek to justify that availability.

Application of the defence

Authority in favour of the defence's availability is admittedly flimsy. *Lipkin Gorman* suggested that it is 'commonly accepted' that a wrongdoer should not be able to avail himself of the defence.[131] Lord Goff did not discuss it further since the question did not arise for decision, although it is hard to think, given the way he made no further comment, that he disagreed with the statement. Henderson J in *Test Claimants in the FII Litigation v HMRC*[132] considered that the wrongdoer bar applied to cases where the defendant is sued for a legal wrong. The common law has, however, diverged with some jurisdictions, such as Hong Kong, taking a very hard line against the application of the defence[133] to wrongdoers and in cases of illegality, and others, like Singapore, being much more liberal.

The standard-bearer case for the availability of the defence in trespass – and by extension other innocent wrongdoing – cases is *Cavenagh Investments Pte Ltd v Rajiv Kaushik*.[134] The decision involved a condominium development at Pebble Bay in Singapore. An employee of the management company forged signatures on the lease agreement, allowing him to lease the apartment to the defendant without the defendant realising that the lease was unauthorised, and he was paying rent to the employee personally. When this came to light

129 Craig Rotherham, 'Morally blameless wrongdoers and the change of position defence' (2018) 30 Singapore Academy of Law Journal 149, 150.

130 See eg Burrows (n 11 above) 647–654; James Edelman, *Gain-Based Damages* (Hart 2002) ch 2.

131 [1991] 2 AC 548 (HL) 580.

132 [2008] EWHC 2893, [320].

133 See *Arrow ECS Norway v AM Yang Trading Ltd* [2018] HKCFI 975, [2018] 5 HKC 317; *DBS Bank v Pan Jing* [2020] HKCFI 368; see, for discussion, Connie H Y Lee and Joshua Yeung, 'Unjust enrichment and illegality: "innocent" wrongdoing and its implications for the change of position defence' [2021] Lloyd's Maritime and Commercial Law Quarterly 51.

134 [2013] SGHC 45, [2013] 2 SLR 543.

the plaintiff sued for trespass and claimed for use-damages. Kaushik claimed change of position and succeeded. Chan Seng Onn J said:[135]

> I do not take the view that there should be a blanket ban on the defence of change of position applying to all cases of restitution for wrongs … Where there is no moral turpitude but the wrong involved is one where the law has prescribed the remedy for a particular policy reason, the defence should also not apply … In the present case I do not see why the defence should not apply

In essence the judge said that the policy behind rendering the conduct wrongful would not be defeated by providing a change of position defence. In making this decision, the judge relied in part on a *dictum* of Lord Nicholls in *Kuwait Airways v Iraqi Airways (Nos 4 & 5)*.[136] That dispute arose from the Gulf War and the conversion by Iraqi Airways of planes taken by Iraqi forces following Iraq's annexation of Kuwait. Lord Nicholls' *dictum* is not completely clear as he apparently believed a claim in conversion to be one in unjust enrichment, saying:

> Vindication of a plaintiff's proprietary interests requires that, in general, all those who convert his goods should be accountable for *benefits* they receive. They must make restitution to the extent they are unjustly enriched. The goods are his, and he is entitled to reclaim them and any benefits others have derived from them. Liability in this regard should be strict subject to defences available to restitutionary claims such as change of position[137]

Conversion is not the same as unjust enrichment, as Chan Seng Onn J recognised;[138] however, the point is that the defence can operate if it is consistent with the policy behind the wrongfulness.[139] Specifically, the policy behind conversion does not require that the innocent defendant, who may not have realised he was interfering with another's rights, suffer the loss consequential on his change of position in reliance.

---

135 Ibid [64]–[65].

136 [2002] UKHL 19, [2002] 2 AC 883 (HL); more recently Henderson J at first instance in *Test Claimants in the FII Litigation v HMRC* [2008] EWHC 2398, [320], suggested that the wrongdoing Lord Goff had had in mind on the facts of *Lipkin Gorman* was conversion.

137 [2002] UKHL 19, [2002] 2 AC 883, [79].

138 [2013] SGHC 45, [2013] 2 SLR 543.

139 *Rose v AIB Group* [2003] EWHC 1737, [2003] 1 WLR 2791; Duncan Sheehan, 'Change of position in insolvency' [2004] Cambridge Law Journal 41; *Skandinaviska Enskilda Banken v Conway* [2019] UKPC 36, [2020] AC 1111, 1153; *Test Claimants in the FII Litigation v HMRC* [2014] EWHC 4302 (Ch) [315] (Henderson J); it seems a stretch to call the Revenue a wrongdoer, but certainly the policy underlying *Woolwich v IRC* [1993] AC 70 is inconsistent with change of position. See also Bant, 'Change of position' (n 37 above) 131.

Justification for the defence

Rotherham has put forward a strong defence of change of position in this context. His first point links the availability of the defence and the rationale for the availability of use-damages. Rotherham argues that the salience of allowing the defence militates in favour of seeing the relief as restitutionary. This is slightly shaky and backwards, but he does provide a more positive case for the defence, noting that the justification for imposing gain-based liability on an innocent defendant is itself shaky at best.[140] Gain-based relief is frequently seen by courts as exceptional, and Rotherham believes there is little merit in rendering liability for gains strict just because liability for losses are strict. The argument is that innocent wrongdoers are in the same moral position as the unjust enrichment defendant.[141] Take, for example, a mistaken payee and one who buys a chattel from a converter. Both are completely unknowing and morally innocent. From here we can conclude that an innocent wrongdoer who exercises her autonomy and pays away the value of property that she has innocently converted should be able to put the risk of that on the claimant and take advantage of the defence. Douglas has also argued that the importance of the claimant's property rights does not in itself justify strict liability to repay all benefits,[142] and this also lies behind Lord Nicholls' suggestion that converters be able to take advantage of change of position.[143] The importance of the property right needs to be balanced against the defendant's freedom of action. Importantly the defendant, if he has changed his position, would not be free to determine his own spending priorities if liability were imposed; if he is an innocent defendant this is unfair.[144] Theft cases where the innocently converted asset is then stolen should probably also count and for the same reason as in unjust enrichment, namely that the innocent (and possibly unknowing) wrongdoer would be forced to introduce unwanted precautions against loss.[145]

The basis of the defence is therefore that good faith defendants should have their freedom of action protected as part of an internal trade-off with the strictness of the liability. As Bant suggests, restitution does not aim to impose loss on a defendant.[146] Subject to overriding

140 Rotherham (n 129 above) 165–166.

141 Paul A Walker, 'Change of position and restitution for wrongs: ne'er the twain shall meet' (2009) 33 Melbourne University Law Review 235, 251–252.

142 Simon Douglas, 'The nature of conversion' [2009] Cambridge Law Journal 198, 220.

143 *Kuwait Airways v Iraqi Airways* [2002] UKHL 19, [2002] 2 AC 883, [79] (Lord Nicholls).

144 Rotherham (n 129 above) 167.

145 Birks (n 10 above) 38.

146 Bant, *Change of Position Defence* (n 37 above) 171.

policy considerations, there is therefore no principled reason why change of position should not apply to strict liability wrongs or claims to vindicate a continuing proprietary right (which conversion arguably does). This is a rather more nuanced position than that of Burrows[147] to the effect that change of position can never outweigh the policy behind wrongdoing. Importantly, the claimant's right to a loss-based remedy remains unaffected by change of position and so this only affects recovery in cases where the defendant's gain is greater than the claimant's loss and reflects the point that anything the claimant recovers over and above her losses amounts to a windfall.

In *Ministry of Defence v Ashman*[148] Mrs Ashman remained in Ministry of Defence (MoD) accommodation after her entitlement to do so ended. The MoD sought to recover mesne profits from her and succeeded. However, although Mrs Ashman had made a deliberate decision to remain in the property, the Court of Appeal did not award the objective market rent as mesne profits, but the lower discounted rate applicable to the type of local authority housing she would have gone into had it been available, which it had not been (at least until the eviction order was made). Although not couched in terms of change of position as such, the reasoning given by Hoffmann LJ for the reduction, which he termed an example of subjective devaluation,[149] was in terms of her having no practical choice but to remain. She was innocent, and the reduction in quantum is consistent with the policy behind trespass. This can legitimately be seen as taking her autonomy into account.[150] We have seen that a link has been drawn between subjective devaluation and change of position via this respect for the defendant's autonomy, and so this provides further support by analogy for using the defendant's autonomy as the foundation for the defence of change of position in these cases of innocent wrongdoing.

Importantly therefore, the justification for the defence in the wrongs context is defendant-sided in that it concentrates on the moral position of the defendant, not that of the claimant. If the 'outcome responsibility' justification for the defence of change of position in unjust enrichment is sound, we must conclude that any justified change of position defence in wrongs cases is not the same defence. An outcome responsibility

147 Burrows (n 11 above) 699–700; see also Ross Grantham and Charles Rickett, *Enrichment and Restitution in New Zealand* (Hart 2000) 354.

148 (1993) 66 P&CR 195.

149 Ibid 201–202.

150 Craig Rotherham, 'Subjective valuation of enrichment in restitution for wrongs' [2017] Lloyd's Maritime and Commercial Law Quarterly 412, 419–422; *Cavenagh Investments Pte Ltd v Kaushik* [2013] SGHC 45, [2013] 2 SLR 543 also accepts subjective devaluation within trespass. See Rachel Leow, 'Change of position in restitution for wrongs: a view from Singapore' (2014) 130 Law Quarterly Review 18.

analysis after all would presumably fix the wrongdoer with the greater level of responsibility than the (also) innocent claimant. This in turn entails that change of position should be inapplicable. Accepting of course that there is a greater weight of authority for denying the defence here, we can, however, say that, if the innocent defendant is made to repay or give up his gains to the claimant, he will be made irreversibly worse off. In *Cavenagh Investments*, Kaushik would in effect be forced to pay rent twice were change of position not available. Since litigation would presumably be needed to recover the payments to the fraudster, those payments would be irreversible in the sense used in *AFSL*. Kaushik would be made worse off, and that is not the purpose of a restitutionary claim.

Extending the defence to proprietary cases

This section is divided into two subsections. First, we examine how the defence works in the proprietary context and see that the mechanics are different from those in personal claims. Secondly, we examine the extent to which the claimant can be said to have outcome responsibility for the defendant's actions.

There are two preliminary matters. First proprietary claims are distinctively different from personal claims. One difference is simply that the mechanics of the defence's operation are different. There is also a question whether change of position is inconsistent with vested rights.[151] If the defendant is a trustee for the claimant, it is difficult to argue that change of position lies without unduly weakening the protection of beneficiaries against trustees.[152] A second difference is that the defendant might not be enriched in the same type of way as in personal claims. First Chambers[153] and later Lodder[154] have claimed that there are two distinct types of enrichment. The defendant might be enriched by value or by rights. Value in this context is relational exchange value, as opposed to simple aesthetic value, and refers to value defined by relation to, reference to and ultimately in exchange for another item. By contrast, where a party is enriched by rights and is so unjustly, the claimant is able to recover that specific right through

151 Elise Bant and James Edelman, *Unjust Enrichment* 2nd edn (Hart 2016) 354; See also *Armstrong DLW GmbH v Winnington Networks Ltd* [2012] EWHC 10; [2013] Ch 156, [103] (Morris QC).

152 On which protection, see Duncan Sheehan, 'Equitable remedies for breach of trust' in Roger Halson and David Campbell (eds), *Research Handbook on Remedies* (Edward Elgar 2019) 146.

153 Robert Chambers, 'Two kinds of enrichment' in Robert Chambers, Charles Mitchell and James Penner (eds), *Philosophical Foundations of Unjust Enrichment Law* (Oxford University Press 2014) 242.

154 Andrew Lodder, *Enrichment in the Law of Restitution of Unjust Enrichment* (Hart 2016) 38–43.

a power to re-vest it,[155] and it does not matter if we think the right valueless.[156] If we accept this dichotomy between enrichment by rights and by value and also that there should be symmetry between disenrichment and enrichment,[157] disenrichment by value should not affect enrichment by rights and the defence of change of position in the proprietary claims' context – if it applies – cannot be based on disenrichment. However, such a blanket ban can be avoided if we adopt the irreversible detriment approach.[158]

The second preliminary point is that change of position might be relevant to subrogation claims,[159] but they do not require explanation separate from the personal claim. In *Boscawen v Bajwa*,[160] the Abbey National's money was held on trust for the purchase of a property owned by Bajwa. That property was subject to a mortgage in favour of the Halifax Building Society. The money was used to discharge the latter mortgage, but no purchase went through and Abbey sought to be subrogated to the Halifax's mortgage. Bajwa would have been enriched by having the debt discharged, was enriched at the expense of Abbey and the money was paid without authority. The mortgage secured a personal debt and the personal unjust enrichment claim will be susceptible to change of position.[161]

Operation of the defence by counter-restitution

It would be wrong if an express trustee having enriched himself (even if innocently) through a breach of trust could defend himself with change of position. Some authors have chosen to distinguish therefore between unexercised powers where the defendant can rely on change of position and the power once exercised after which the defendant cannot.[162] Rescission claims, for example, are often, although not universally, held up as involving a power. Birke Häcker is often seen as the leading proponent of this idea.[163] Häcker distinguishes between the

155 Ibid 64–65

156 Chambers (n 153 above) 242–243, discussing the decision in *Blacklocks v JB Developments (Godalming) Ltd* [1982] Ch 183. Lodder (n 154 above) 112; *Armstrong v Jackson* [1917] 2 KB 722; *Balfour & Clarke v Hollandia Ravensthorpe* (1978) 18 SASR 240.

157 On which see Liu (n 4 above) 306.

158 Bant, *Change of Position Defence* (n 37 above) 206.

159 *Boscawen v Bajwa* [1996] 1 WLR 328 (CA) 341; *Goff and Jones* (n 5 above) para 27.70.

160 [1996] 1 WLR 328.

161 Ibid 340–341; *Filby v Mortgage Express Ltd* [2004] EWCA Civ 759.

162 See eg Ben McFarlane, *The Structure of Property Law* (Hart 2008) 333–335; Bant, *Change of Position Defence* (n 37 above) 206–207.

163 Birke Häcker, 'Proprietary restitution in impaired consent transfers: a generalised power model' [2009] Cambridge Law Journal 324.

'immediate interest' model, generating a trust and the 'power model' arguing that both in common law and equitable rescission the claimant has a power *in rem*.[164] In the tracing context, the ability of the trust beneficiary (say) to claim against a third party is also, controversially, said to be based on a power.[165] This is not universally accepted either, although, if we do accept it, there is an analogy with rescission where change of position can definitely be taken into account.[166] Penner, however, argues that the beneficiary is simply electing to enforce the trust interest against the third party and that this is a feature of the beneficiary's interest in a trust fund, and not an interest in particular assets.[167] This analysis need not preclude change of position though. It is possible to accept the operation of change of position in proprietary cases while subscribing to the interest in a fund analysis.[168]

Where the claim operates by means of a power, the operation of the defence is by means of a counter-restitutionary payment,[169] or by way of set-off.[170] If the claimant has a tracing claim over a painting in the hands of the defendant, who had saved £150 to buy a new picture but now spends it on a celebration dinner, the claimant can only assert the equitable right if she pays £150 to the defendant.[171] The traceably surviving right should only be recoverable if the claimant is prepared to give credit for the change of the defendant's position. It is important that the money paid away or spent comes from a source unconnected traceably to the assets over which the claim is made. If the money paid away is traceably derived from the initial receipt, change of position is unnecessary as the traceable assets recoverable have reduced. The counter-restitution requirement is inevitable. If the defendant has a right that is traceably derived from the claimant's right, that right cannot be divided. The defendant either has it or not. The claimant either has a claim or not, and so making restitution conditional on counter-restitution is the only way of implementing the defence.

---

164 Ibid 328–331.

165 Aruna Nair, *Claims to Traceable Proceeds* (Oxford University Press 2017) para 6.42.

166 Bant, *Change of Position Defence* (n 37 above) 207.

167 James Penner, 'The (true) nature of the beneficiary's equitable proprietary interest under a trust' (2014) 27 Canadian Journal of Law and Jurisprudence 473.

168 Peter Jaffey, *Private Law and Property Claims* (Hart 2007) 161–163.

169 Jaffey's model does not operate like this, but there are significant criticisms to his model; see Duncan Sheehan, 'Express trusts, private law theory and legal concepts' (2022) 35 Canadian Journal of Law and Jurisprudence 511–536.

170 *Alati v Kruger* (1955) 94 CLR 216.

171 McFarlane (n 162 above) 334; Häcker (n 163 above) 347–349; *AH Macdonald & Co Pty Ltd v Wells* (1931) 45 CLR 506 provides an example of an innocent misrepresentor taking advantage of change of position.

Application of outcome responsibility to the claim

This requires us to cleanly separate out a number of scenarios. Rescission claims are, albeit controversially, said to be unjust enrichment claims. However, rescission claims rely on mistake, duress or undue influence and calling them unjust enrichment claims draws attention therefore to the unity in the reason for restitution.[172] In the first scenario where the claimant wishes to rescind against the immediate recipient of the asset and recover the very asset transferred, change of position should be available for the same reason as in the personal unjust enrichment cases. If, as a result of a mistake, a deed is voidable, but the beneficiary of that deed has paid away money (from a different account) and a sufficient causal connection can be found between the claimant's mistake and the defendant's payment, the claimant should be seen on Ratan's view as having sufficient outcome responsibility for the payment away. By the same token, the defence will not be available, as we saw earlier, if the defendant is at fault or in bad faith in some way for causing the transfer.[173]

A rescission claim may reach substitute assets through tracing; this is our second scenario. In *Bainbridge v Bainbridge*,[174] Master Matthews commented that rescission founded claims to property other than that initially transferred.[175] Where the substitute property remains in the hands of the initial transferee who then pays away money (again from a different account) in reliance, the defence must continue to apply – and again for the same reason as in personal unjust enrichment claims.

The third scenario is where the claimant attempts to claim against a third-party donee. Such parties are also vulnerable to rescission,[176] and claims by a beneficiary of a trust may likewise extend to remote transferees. Some transferees with notice are vulnerable to rescission,[177] although the presence of notice makes it harder to see the availability of change of position. It is controversial whether these are unjust enrichment claims. *Foskett v McKeown*[178] is authority that they are not. Birks, however, argued strongly that they are.[179]

172 Bant, *Change of Position Defence* (n 37 above) 90–91.

173 See, in this context, Robert Chambers, 'Proprietary restitution and change of position' in Dyson et al (eds) (n 7 above) 115, 123.

174 [2016] EWHC 898.

175 Ibid [24]–[32]; see also Dominic O'Sullivan, Steven Elliott and Rafal Zakrzewski (eds), *The Law of Rescission* 2nd edn (Oxford University Press 2014) paras 21-04–21.05.

176 *Bainbridge v Bainbridge* [2016] EWHC 898, [24]–[32]; *Load v Green* (1846) 15 M&W 216, 153 ER 828.

177 O'Sullivan et al (n 175 above) paras 21.08–21.21.

178 [2001] 1 AC 102 (HL).

179 Birks (n 4 above) 198.

Whether we think the claim lies in unjust enrichment or not, the power of Ratan's argument reduces as the claimant traces into the hands of third parties and there is, as in *Foskett v McKeown*, a tracing but no causation link.[180] The additional steps in the chain make it ever harder to say that the actions of the claimant led to the defendant's reliance. In cases like *Foskett* the claimant trust beneficiary has no real non-fictional responsibility for the asset transfer by the trustee at all; it is impossible to see how he has, in fact, contributed to the outcome. The House of Lords, in fact, not only rejected the idea that these are unjust enrichment claims but also rejected change of position. This supports Chambers in his claim that, even if the proprietary claim against the initial recipient is sourced in unjust enrichment and is susceptible to the defence, a claim against subsequent transferees being sourced in property is not sourced in unjust enrichment. The property right is enforced because it is a property right. This has an important corollary on Chambers' view; if unjust enrichment is irrelevant, so is the defence of change of position.[181]

However, the strictness of the liability of the innocent third-party donee to return assets and the effect of the claim on his creditors in insolvency can be set against the need for the donee's freedom of action to be respected. That the defendant (and by extension his creditors) knows nothing of, and cannot guard against, the claimant or his claim militates in favour of the defence, and actually militates in its favour irrespective of whether we think the claim is an unjust enrichment claim. The defendant's moral position should not depend on which bank account he decides (arbitrarily) to withdraw from. If the third-party tracing defendant takes money from a separate unconnected account and is forced to repay the defendant, he will be made irreversibly 'worse off' than the *status quo ante*. This is the defendant-sided justification which was so powerful in cases of innocent wrongdoing and which becomes more powerful in the proprietary context as any suggestion of the claimant's outcome responsibility recedes. Ratan's outcome responsibility thesis has no purchase in these cases. The justification in *AFSL* therefore has purchase here too, and, indeed, Bant has shown how this characterisation of the defence – irreversible detriment – ties with the principle of *restitutio in integrum* in rescission.[182] Once we accept that change of position is not tied to an unjust enrichment cause of action and disentangle it from the enrichment inquiry by reference to irreversible detriment, the way is open to accept the defence in

180 Duncan Sheehan, 'Subtractive and wrongful enrichment: identifying gain in the law of restitution' in Charles Rickett (ed), *Justifying Private Law Remedies* (Hart 2008) 331, 346.

181 Chambers (n 173 above) 128–129.

182 Bant, 'Change of position' (n 37 above) 149.

misapplied trust property and rescission cases – even if they are not unjust enrichment claims.[183]

CONCLUSION

The implications of Ratan's view of change of position in unjust enrichment is that the defence can be fitted into a framework whereby the two parties – claimant and defendant – are locked together through the medium of the claimant's outcome responsibility for the defendant's position. Change of position, justified in this way, might encompass both cases of personal claims and at least some proprietary claims, particularly rescission or subrogation claims, although it struggles to accommodate the unjust enrichment defendant's independent changes of position such as the object's theft; such might need a separate explanation. There is authority that change of position may also stretch to cases of innocent wrongdoers. There are good normative reasons why such parties should be granted a defence, although the outcome responsibility argument in personal unjust enrichment claims has no purchase. Nonetheless, internal trade-offs between the defendant's freedom of action and the strictness of liability in these torts render the defence justifiable, and this rationale also has purchase in the context of proprietary claims against remote recipients, where again the outcome responsibility justification seems to have little, if any, purchase.

Ratan developed his view in the context of a desire to find a single baseline against which to judge how far the defendant had changed her position in both anticipatory receipt cases and post-receipt reliance cases. We might question whether he really needed to build such an elaborate edifice for such a purpose, particularly given the need to construct an apparatus around assessing causal contribution, and the need to hunt for a different explanation in those cases of wrongs, independent changes of position, and proprietary claims where the application of change of position seems ethically defensible, and in wrongs cases supported by some authority, but where outcome responsibility has no purchase. It is true, of course, that in assessing 'detriment' some baseline – preferably single baseline – needs to be picked and there remains a difficult question as to how to justify the baseline. In other words, Bant's thesis allows us to justify change of position outside personal unjust enrichment claims – in some restitution for wrongs cases and some proprietary claims – but it assumes rather than fully explaining why the defendant should not be made worse off. Ratan's justification does not and cannot, however, fully fill the gap in Bant's thesis that needs filling. Either we cut back

183 Bant, *Change of Position Defence* (n 37 above) 208–209.

the defence substantially, or we continue to work to justify why the defendant should not be made worse off. It is submitted that to cut back the scope of the defence so substantially would be a retrograde step; defendants who deserve to be exculpated from liability would be drawn back into the ambit of liability. More work is therefore needed.

Northern Ireland
Legal Quarterly
Summer Vol. 74.2 (2023) 303–322
Articles DOI: 10.53386/nilq.v74i2.1097

Property guardians: reigniting the lease/licence battle?

Dean Taylor*

Southampton Law School, University of Southampton
Correspondence email: dean.taylor@soton.ac.uk

ABSTRACT

This article analyses recent English decisions reviving the need to consider the lease/licence dichotomy and conclusiveness of the parties' agreement in the new context of property guardianship as an alternative to private renting. It argues that context has proved instructive in interpreting the parties' agreement elsewhere in the case law and offers a way forward in the hard cases amid the ongoing search for doctrinal clarity and justification. A compound subjective–objective approach appreciates the underlying purpose of the parties' relationship and justifies why no intention to grant the right of exclusive possession can be present, thereby precluding a tenancy. The article briefly considers reforms to rental accommodation previously suggested by the Law Commission and, in light of the continued need to prove the status of lessee, argues that they should be revisited in order to protect those living in temporary accommodation.

Keywords: property guardians; property; housing; leases; licences; exclusive possession; intention; interpretation.

INTRODUCTION

In recent years, the English County Court[1] and High Court[2] have respectively considered the legal status of property guardians. The cases illustrate how property guardianship has now become mainstream as an alternative form of urban, 'meanwhile' housing/living. Despite being low-level decisions, *Roynon* and *Khoo* are both of significance as they offered the first substantive opportunity for the courts to consider property guardians' legal status amid a changed housing landscape – guardians having until now existed in the grey area between leases and licences. However, that the cases reached contrasting conclusions means clarity remains elusive.

* Lecturer in Law. Many thanks to the anonymous reviewer for their comments, and, too, to Professor Sue Bright and Dr Craig Prescott for theirs on an earlier draft. Any *errata* are the responsibility of the author.

1 *Camelot Property Management Ltd v Roynon* (Bristol County Court, 24 February 2017) (*Roynon*).

2 *Camelot Guardian Management Ltd v Khoo* [2018] EWHC 2296 (QBD) (*Khoo*).

Given the increasing use of property guardians as a housing alternative, there is a pressing need to resolve their status. The struggle in the relationship between landlord and tenant is being revived – a struggle thought to have been settled following *Street v Mountford*[3] where a true construction of the agreement coupled with exclusive possession, for a term certain (or periodic), and (usually) at a rent reveals whether a tenancy exists.[4] However, these requirements require qualifying by reference to what else was said in the case, as well as surrounding authorities. The qualification is particularly acute when resolving the tension between finding exclusive possession while also giving effect to the parties' agreement in the way they intended.[5] Ockham's razor presents itself as the court is required to explain the simultaneous veracity of the parties not appending their own label to the agreement while at the same time, where there is exclusive possession, there can be no tenancy if the parties do not intend to create one.[6] The legacy-factors of vigilance against sham transactions in the residential context find their expression in *Roynon* and *Khoo* respectively, if inconsistently, and are sensitive, given that property guardianship lies between providing temporary accommodation in fulfilment of a wider commercial purpose. The hybrid nature of property guardianship requires a careful assessment of the application of the *Street* criteria as further coloured by *Bruton v London & Quadrant Housing Trust* where the non-estate-owning intermediary trust was still held to have conferred a tenancy.[7] Similarly, in the property guardianship scheme, the guardian company often has not been granted an estate in the land in its own right and, relative to the proprietor, a guardian *can* be a tenant for the purposes of the legislation.[8] The court is caught between an interventionist pursuit of achieving long-standing social policy/distributive justice goals versus regulation of a private, consensual and contractual bargain. In addition to this are the contentions between effect and substance; distinguishing between what is genuine from what/when it is not; and the struggle between pragmatism and principle. Common to both *Roynon* and *Khoo* was a less than full consideration of the existing jurisprudence beyond *Street* and *Bruton*. This article offers a way forward by situating property guardianship within the existing purpose-driven approach found in other contexts

3 [1985] AC 809 (HL).

4 Ibid 818C, 818E, 826G, 827A–B (Lord Templeman) .

5 *Newham LBC v Hawkins* [2005] EWCA Civ 451 (CA), [36] (Arden LJ).

6 S Bright, 'Street v Mountford revisited' in S Bright (ed), *Landlord and Tenant Law Past, Present and Future* (Hart 2006) 21.

7 [1999] UKHL 26, [2000] 1 AC 406.

8 Housing Act 1988, s 21; Protection from Eviction Act 1977, s 3 provides minimum notice period for licensees.

within private law, where the words of the parties' agreement and their effect coincide in the context of fulfilling a genuine (outer) purpose. The article argues that the impact of the language of sham/pretence has been blunted in the aftermath of *Street*, and clearer articulation of the categories of cases is needed which in certain circumstances aligns with the parties' written agreement. The article acknowledges the limitations and drawbacks of adopting a *laissez-faire* approach; the coda suggests legislative reform has now become more urgent to maintain realising housing law's aims amid an increasing shift towards formalism.

CONTEXT

The premise of property guardianship is simple: property becomes vacant (often subject to seeking planning permission and (re)development), thus risking vandals/squatters and the concomitant expense of eviction proceedings. Instead, the owner licences day-to-day management to a guardian company which in turn installs individuals to occupy and guard the property 'round-the-clock'. Property guardianship originates from the Netherlands where it is an accepted alternative to traditional renting. With a number of central 'players' in the sector, the scheme has gained traction in England since the early-2000s as an emergent form of insecure low-cost housing.[9] Property guardianship has attracted frissons of media excitement as a solution to the desperate shortage of affordable housing, providing security for proprietors in urban contexts. The benefits of the scheme centre upon the relatively inexpensive rents payable, comparably larger living space, and greater autonomy and flexibility compared to traditional private rental.[10] Despite these benefits, recent scholarship has located property guardianship within the theoretical framework of precarity.[11] The transitory nature of dwelling that property guardianship entails is suffused with a temporal precarity and the elusive promise of security – in the narrow and wide senses of tenure and affectively in respect of place in society. As Ferreri et al elucidate, precarity is embodied within property law: its etymology deriving from *precarius*, referring

9 M Ferreri, G Dawson and A Vasudevan, 'Living precariously: property guardianship and the flexible city' (2016) 42 Transactions of the Institute of British Geographers 246–259.

10 C Clemoes, 'Property Guardians, London: The State of Housing in Transient Places' (MSc thesis 2014) (on file with author) 24–25 and 53.

11 C Hunter and J Meers, 'The "affordable alternative to renting": property guardians and legal dimensions of housing precariousness' in H Carr, B Edgeworth and C Hunter (eds), *Law and the Precarious Home: Socio-Legal Perspectives on the Home in Insecure Times* (Hart Publishing 2018) 65–86.

to property 'held by tenancy at will, uncertain, doubtful, suppliant'.[12] That precarity's origins are spatialised and operationalised in relation to a tenancy coincides with the courts' attention on the legal status entailed by the scheme, and which needs resolving.

ROYNON AND *KHOO*

The Facts

The facts in both cases can be briefly put. The contention made by each guardian was that the guardian company could not evict them in reliance on the notice period set out in the respective agreements.

Roynon

Camelot Property Management Ltd (CPML) and Camelot Guardian Management Group Ltd (CGML) sought possession of the Broomhill Elderly Persons Home in Bristol. Bristol City Council engaged Camelot in 2013 to install guardians, and in January 2014 the defendant moved in. A notice to quit was served in May 2016 and, after refusing to leave, possession proceedings were initiated. In order to determine whether the court could grant an order for possession, it needed to be ascertained – as a preliminary issue – whether Mr Roynon was a licensee or an assured shorthold tenant, and whether the lease had in fact been determined.

Khoo

Khoo concerned an appeal against the finding that the defendant was a licensee. CGML having sought possession of the property in Soho, central London, claimed the guardian was a trespasser. At first instance it was found that *de facto* exclusive possession was enjoyed – however, it was also found that the occupation of the property was not a tenancy, but a licence as described.[13] In August 2017, the owners, Westminster City Council, gave notice it would need the property back in order to begin redevelopment. One month's notice was given that the licence would end on 11 October 2017, but Mr Khoo remained in occupation as the sole occupant at the time possession proceedings ensued. It was not disputed that if a tenancy arose it was an assured shorthold tenancy (AST), and that if it was found to be an AST it had not been determined and the claim for possession should be dismissed.

---

12 Ferreri et al (n 9 above) 249.

13 *Khoo* (n 2 above) [1].

The decisions

To understand the decisions is to start with the occupation agreement where the labels employed can obscure the substance of the relationship. In recitals to the standard-form terms and conditions, the following is provided:

> 1.3 Camelot provides services to property owners to secure premises against trespassers and protect them from damage (among other things) and has agreed to provide such services to the Owner in respect of the Property
>
> 1.4 *To enable Camelot to provide those services* the Owner has agreed that during the period permitted by Camelot's agreement with the Owner *Camelot shall be entitled to grant temporary non-exclusive licences to share occupation of the Property which do not confer any right to the exclusive possession* of the Property or any part of it
>
> 1.5 *Camelot is not entitled to grant possession or exclusive occupation*[14] of the Property or any part of it to any other person. It merely has the power to grant licences for non-exclusive occupation of the Property.[15]

Further,

> The parties agree that this agreement is not intended to confer exclusive possession upon the Guardian nor to create the relationship of landlord and tenant between the parties and the Guardian shall not be entitled to an assured tenancy or a statutory periodic tenancy under the Residential Tenancies Act 2004 or to any other statutory security of tenure now or upon the determination of the Licence.[16]

---

14 It is not immediately clear why 'exclusive' is used to describe the *occupation*, but not possession. It does not follow that exclusive possession and exclusive occupation are the same thing: *Westminster Council v Southern Railway* [1936] AC 511.

15 Emphasis added. NB the agreements were available on CPML's website and were accessible/accessed by the author on 16 November 2018, but now non-extant. A search at Companies House indicates resolutions were passed to wind up CPML and CGML in April 2017 and November 2019 respectively.

16 Emphasis added. It is not particularly clear why an Irish Act is referred to here given the disputes' locus was England. No reference is made to this in either case, but were this to have been addressed, it might give credence to the pretence argument: a term inserted with no intention of being operated by (one of) the parties (*Street v Mountford* (n 3 above) 825H; *AG Securities v Vaughan; Antoniades v Villiers* [1990] 1 AC 417, 462H (Lord Templeman), 470A (Lord Oliver)).

Roynon

Whilst headed 'licence', the agreement nevertheless created a tenancy. The agreement stated that Camelot may 'from time to time *designate* [parts of the Property] as being available for *shared* residential use of the Guardian' and referred to the Guardian '*shar*[*ing*] the occupation' and not conferring 'a right to use any specific room'.[17] The reality was different from that envisioned, however. Guardians had a choice of room to occupy and no other guardian would have keys to access the chosen room. However, while the aim and reality did not marry, these of themselves did not make exclusive possession axiomatic.[18] The relationship encompassed a reading of the agreement which contained a number of obligations owed by the guardian (no smoking; no overnight guests; no inviting more than two guests at any one time; no leaving of guests unsupervised). Camelot sought to rely on this as showing no right to exclusive possession and no ability to exclude the world at large from the property.[19] While these obligations placed 'significant limitation' and 'onerous restriction[s]' upon the guardian, they were not incompatible with exclusive possession.[20]

HHJ Ambrose refused to accept the argument that because there was no express reservation of a right to inspect the property, this was indicative of a licence and did not satisfy exclusive possession.[21] The judge reasoned that, while a tenancy agreement may typically contain an express right for the landlord's limited entry to inspect the property, it does not mean that in the absence thereof the arrangement is to be viewed axiomatically as a licence.[22] That Camelot carried out these inspections was also not incompatible with exclusive possession. Therefore, the guardian did have exclusive possession and enjoyed a monthly, periodic AST.[23]

Khoo

Butcher J held that the agreement did properly constitute a licence. Construing the agreement as one referring to a right in respect of the property 'as a whole, not a room or other part of the Property' meant, on a natural meaning of the words, that the occupation could not entail a right of exclusive possession of any part of the property.[24] This

17 *Roynon* (n 1 above) [30], emphasis added.
18 Ibid [33], [36].
19 Cf W*estminster City Council v Clarke* [1992] 2 AC 288 (HL).
20 *Roynon* (n 1 above)[39], [41].
21 Ibid [44].
22 Cf *Street v Mountford* (n 3 above) 818C (Lord Templeman).
23 *Roynon* (n 1 above) [48]–[52].
24 *Khoo* (n 2 above) [21]–[22].

flowed from the series of clauses referring to shared occupation with other guardians, the personal nature of the occupation, and CGML's reserved right to alter the extent and location of the living space. Some disquisition followed as to the precise nature of the obligations guardians owed, and it was concluded that these would be so 'unusual' as to be inconsistent with enjoying exclusive possession over the property to have a tenancy.[25] Where there was a provision that there would always be enough living space for at least one room each, this was only a management device to prevent introducing more guardians than there were rooms rather than providing a room exclusively for each guardian.[26]

Butcher J also found that the way in which the room was 'let' did not raise enough significance as to 'relevant context' to constitute exclusive possession for the purposes of a tenancy. Just because it was envisaged that a particular room was to be made available, it did not follow that the terms of the agreement should be construed in a way the language itself did not point toward. The parties' agreement had to be read holistically in light of the scheme *per se* which maintained from the outset that it did not lie in the gift of Camelot to grant exclusive possession over any part of the property – itself only holding a licence. The agreement describing the parties' relationship was a true bargain not containing any element of sham or pretence, which ordinarily connote elements of dishonesty.[27] While at first instance the agreement was inferred as 'almost certainly intentionally misleading' on the part of CGML, this was an inference that could not be safely drawn to overcome the presumption that the parties intended their agreement to be taken at face value.[28]

DISCUSSION

While property guardianship's existence has been 'pre-legal', it has now been thrust into the classic lease/licence dichotomy. *Roynon* and *Khoo* speak to the characterisation of the 'rough-and-ready grasp of the empirical realities of life' and conceptual model of property as fact.[29] In the property guardianship context, this is reflected in English law's primitive focus on possession. However, the inconsistencies between *Roynon* and *Khoo* speak to a lack of doctrinal clarity, making

25 Ibid [23].
26 Ibid [24].
27 Ibid [19].
28 Ibid [37]–[38].
29 K Gray and S F Gray, 'The idea of property in land' in S Bright and J Dewar (eds), *Land Law: Themes and Perspectives* (Oxford University Press 1998) 15, 18–20.

it practically difficult to pinpoint the court's foci when ascertaining the proper nature of the parties' relationship.

Exclusive possession

Roynon and *Khoo* remind us that, while the parties may describe their agreement as a licence, this description cannot alter its substance. Their diametrically opposed conclusions illustrate the continuing elasticity in interpreting exclusive possession and centres particularly upon justifying those circumstances where it is appropriate to deny the right.[30] The undercooked analysis of what exclusive possession means is not new, and principled clarification is needed. What lacks in linguistic deftness has some metaphorical truth where Gray and Gray refer to the schizophrenic approach of property law's logic,[31] and finds its expression in possession's meaning oscillating between question of fact and law.[32] In the lease/licence context allowing the court to use subsequent conduct to aid its construction of the agreement, this deems the right to have arisen by fact, rather than lying in grant.[33] Such an outlook is described by Crawford as reflecting an expressive theory of possession, justified from observing the interaction of people with things as a 'social fact'.[34] However, notwithstanding this attempt to theorise possession as essentially fact-based, it still remains that possession is rights-based too.[35]

That guardians are exhorted to treat the property 'as if ... their own' connotes the air of control and exclusory power, but this is to confuse a behaviourist attitude towards exclusive possession such that it is being strained to fit the doctrinal requirements. While uninterrupted enjoyment can be evidence of the fact of exclusive possession (as in *Roynon*), this commits to the tautology of exclusive possession arising simply through an exercise that often conflates occupation

30 *Street v Mountford* (n 3 above) 817E, 817G, 826G–827B.

31 K Gray and S F Gray, 'The rhetoric of realty' in J Getzler (ed), *Rationalizing Property, Equity and Trusts: Essays in Honour of Edward Burn* (Butterworths 2003) 204, 221.

32 *Bruton v London & Quadrant Housing Trust* (n 7 above) 413G–H (Lord Hoffmann). In *Khoo* (n 2 above) reference is made to the first instance judge's finding that *de facto* exclusive possession and the right thereof was enjoyed (at [2]).

33 *Street v Mountford* (n 3 above), 819E–F cf *Onyx v Beard* [1996] EGCS 55 (Ch) (HHJ Hart QC): 'The correct characterisation of the relationship is, of course, entirely independent of the results which flow from it.' Further, discussion in *AG Securities v Vaughan; Antoniades v Villiers* (n 16 above) 464A (Lord Templeman), 469C (Lord Oliver).

34 M Crawford, *An Expressive Theory of Possession* (Hart 2020) 7.

35 S Green and J Randall, *The Tort of Conversion* (Hart 2009) 86.

with possession.[36] If factual exclusive possession were the only determinant, then this risks assuming the thing that is being proved. The conception of possession across property law requires not only factual possession but also the requisite intention to possess (*animus possidendi*).[37] Goymour suggests the *Bruton*-tenancy (present in *Roynon*) is an example of original, relative title arising from the fact of possession.[38] Recent cases follow in the vein of *Bruton* where original, relative title through the fact of possession is enough to defend an application for rectification of the Register.[39] However, while there is utility in this approach, it is not justified in all cases. From an orthodox perspective, the derivative nature of the lease speaks to the consensual dependency upon which possession exists in this *particular* context and where property guardianship is not quite of the 'self-help' ilk as adverse possession. Thus, *de facto* possession should only be a step along the way of assessing whether a tenancy exists rather than being the inexorable touchstone some cases suggest it is. What is needed is a regularising of the dependency upon which exclusive possession hangs in a tenancy and which *Khoo* seeks to restore. Even if there is occupation of land, this is not synonymous with its control; and, if such aspects of control associated with ownership are absent as a matter of fact, then this should be influential that no right has been granted.

Street avers that where the *only* circumstances are that residential accommodation is offered then a tenancy arises.[40] In *Street*, Lord Templeman accepts that the statute is 'irrelevant' to determining the effect of the parties' agreement, nor too will it (there the Rent Acts) alter the effect of the agreement.[41] Notwithstanding the canonical quasi-statutory status as a major premise affecting later cases, his Lordship's reasoning suffers from ambiguity. The jurisprudence is replete with minor premises deriving from *Street*, but what of those cases that cannot be straightforwardly described as 'ordinary' residential accommodation?[42] What of those cases where the very essence of

36 P Vincent-Jones, 'Exclusive possession and exclusive control of private rented housing: a socio-legal critique of the lease–licence distinction' (1987) 14(4) Journal of Law and Society 445, 449; S Bright, H Glover and J Prassl, 'Tenancy agreements' in E Simpson and M Stewart (eds), *Sham Transactions* (Oxford University Press 2008) 105, 112.

37 *JA Pye (Oxford) Ltd v Graham* [2003] 1 AC 419, [40] (Lord Browne-Wilkinson), [71] (Lord Hope).

38 A Goymour, 'Bruton v London & Quadrant Housing Trust' [2000]: relativity of title, and the regulation of the "proprietary underworld"' in S Douglas, R Hickey and E Waring (eds), *Landmark Cases in Property Law* (Hart 2017) 161ff.

39 Eg *Rashid v Nasrullah* [2018] EWCA Civ 2685.

40 *Street v Mountford* (n 3 above) 827A–B.

41 Ibid 819G.

42 After *Crancour Ltd v Da Silvaesa* (1986) 18 HLR 265, 280 (Gibson LJ).

the occupation would be undermined?[43] The reasoning in *Roynon* succumbs to this major premise without further analysis. Despite the disconnect between the agreement and reality of the occupation, the latter should not of itself be enough to enjoy the right of exclusive possession, yet there was still a tenancy on the basis the agreement was misleading.[44] The reasoning in *Roynon* enters into a disquisition that narrowly focuses on the clauses concerning the horizontal relationship with fellow guardians, whereas consideration also needed to be given to the vertical relationship vis-à-vis the guardian company. The former would give *carte blanche* to individuals arrogating to themselves control when the sphere of regulation lies properly in a vertical manner. As to the vertical regulation in *Roynon*, the ability of the guardian company to allocate a room is adverted to as conferring exclusive possession of a room, but overlooked (having earlier in the judgment acknowledged clauses elsewhere, albeit – which is also part of the problem – without comment) permissively reserving a power to alter the extent and scope of the accommodation,[45] and can be contrasted directly with the defensible, inductive reading of the living-space alteration clause in *Khoo*, going to the control retained by the guardian company.[46]

If we consider the impact of *Bruton*, the reasoning of Lord Hoffmann there suffers from the same problem of circularity as that in *Street*. While the parties are free to describe their agreement as they wish, meaning and classification are separated in his Lordship's *dicta*, but meaning and classification cannot be easily divorced in the way averred by his Lordship. If the meaning is reflective of the parties' intentions, then so too can (/should) there be accommodation of the effect (barring any sham/pretence) it creates. The unremarkable/superficial elements of control in any agreement will be seen through by the courts,[47] but the nature of the property guardianship scheme should entail a more nuanced consideration, and the reasons for denying control provide a way forward for ascertaining status. As discussed above, the properties are atypical in their accommodation, often not envisaging dwelling for residential purposes. Nevertheless, the ability of the guardian company to install fellow guardians into these atypical buildings takes property guardianship and questions of control beyond previous cases concerning rights of entry/inspection and sharing/installation. This is one reason why the argument that there was exclusive possession of *the whole property* in *Khoo* foundered – if so, only arising by fact (as

43 As discussed in *Khoo* (n 2 above) [26]–[29].
44 *Roynon* (n 1 above) [33].
45 Ibid [30]
46 Ibid [22].
47 Ibid [39].

the last remaining occupant) rather than lying in grant. Whether or not there is a legitimate reason to deny exclusive possession by retaining control through the reserved right to install others to share in the occupation invites rehabilitating *dicta* earlier in *Marchant v Charters* where Lord Denning justifiably sought to explain how various factors will influence the decision: all the circumstances require bearing in mind and thereby avoid ascribing exclusive possession not borne out by the parties' common, substantive intention.[48] As Barak elucidates in the search for the norm within the text of the parties' agreement, there are two meanings present in its interpretation – semantic (subjective) and legal (objective).[49] Wholly subjective then wholly objective methods of interpretation can be traced in the development of the lease/licence jurisprudence, but what is needed is space for a compound, integrated subjective–objective approach that can give effect to common authorial intent where appropriate.[50]

The question remains how room can be made for pragmatism given the way in which the law has become increasingly trammelled. As ever, ascertaining the status of an occupier of land depends on whether an internal route is taken to considering the effect of the agreement on its face or an external route is opted for encompassing subsequent conduct.[51] Here, internal and external can take on a wider meaning to speak to the narrow, internal approach of housing law vis-à-vis the wider, holistic approach that appreciates the overlapping juncture, contexts and axioms (but not necessarily opposition) of property, contract and housing law together, where all bear upon regulating jural relationships respecting land.

Which surrounding circumstances?

Albeit the parties cannot append their own conclusive label, the parties' bargain in the agreement, if informed by genuine surrounding circumstances, should be treated as indicative of the nature of their relationship and that no right of exclusive possession was conferred. The distinction between contractual licence and tenancy that Lord Templeman is concerned not to undermine is itself undermined by *dicta* stating that exclusive possession is not of itself conclusive but depends on a number of fact-specific issues.[52] The difficulty lies, given the narrow articulation in both *Street* and *Bruton*, in knowing which of these surrounding circumstances are genuine enough to displace

48 [1977] 1 WLR 1181, 1185F–G.

49 A Barak, *Purposive Interpretation in Law* (Princeton University Press 2005) 3, 6–7.

50 Ibid 32–33.

51 After *Welsh Development Agency v Export Finance Ltd* [1992] BCLC 148, 186.

52 *Street v Mountford* (n 3 above) 823D.

the (effective) presumption of tenancy and which the courts need to approach inductively. While there is the need to remain astute so as not to make the distinction between lease and licence wholly unidentifiable, it is also necessary to give effect to the purposes for which the parties have mutual understanding and a common intention, and which will not always point in the direction of sham. As recent cases have shown, the textual and contextual dynamics both inform interpreting the nature of the parties' relationship.[53]

Surrounding contextual issues informed the decisions, *inter alia*, in *Hunts Refuse v Norfolk*[54] and *Onyx v Beard*[55] respectively, where the analysis could not be stripped away from context and for which a wide view beyond a surface, factual exclusive possession is required. In *Hunts Refuse*, a clause referring to the grantor's reasonable access for the extraction of minerals was to be read not as the reservation of the right of entry, but as a covenant allowing the grantor to exercise the right. The deed did not circumscribe the grantor's continued right to excavate minerals from the land. In *Onyx*, the purpose of the occupation providing land for a social club meant a tenancy was unsustainable; the desire to make the premises available to staff was incompatible with granting an interest in the land. There was nothing in the surrounding circumstances to displace the centrality of the grant of an exclusive licence only (construing *Street v Mountford* narrowly). It was recognised that the club in *Onyx* enjoyed a conceptually different exclusive occupation over the land which further fed into the consideration that while 'there can be no tenancy without exclusive possession, there may still be a licence even though the licensee enjoys de facto exclusive possession'.[56] References to *de facto* exclusive possession are unhelpful in serialising the exclusive possession criterion when, rather, the consideration is composite.[57] In *National Car Parks Ltd v Trinity Development Company (Banbury) Ltd*,[58] clauses requiring the occupier's reasonable assistance allowing the grantor to resurface the land were to be read as obligations under a covenant.[59] These were included to secure the grantee's co-operation rather than reserving a right of re-entry. This might have been a stretch in construction too far, but the requirement of co-operation was held

53 *Arnold v Britton* [2015] UKSC 36, [14]ff (Lord Neuberger); *Wood v Capita Insurance Service*s [2017] UKSC 24, [13] (Lord Hodge): 'textualism and contextualism are not conflicting paradigms in a battle for exclusive occupation'.

54 [1997] 1 EGLR 16.

55 [1996] EGCS 55.

56 *Woodfall: Landlord and Tenant* 28th edn (Sweet & Maxwell 2019) [1.023].

57 Vincent-Jones (n 36 above) 452.

58 [2001] EWCA Civ 1686.

59 Ibid [35] (Arden LJ).

to amplify the retention of control over the land rather than belonging exclusively to the occupant. While the consideration focuses upon substance not form, it did not follow that what was said is irrelevant to the considerations.[60] While not of itself determinative, how the parties express their agreement is relevant and instructive in determining its substance and should be a starting point for the court's holistic consideration.[61]

These cases are indicative of the difficulties attached to transposing the *Street* criteria to the commercial/quasi-commercial context where the court should not be so overzealous that it reconstructs the bargain into something the parties themselves did not intend.[62] Indeed, because of the ability to obtain legal advice, the proclivity toward finding a tenancy cannot easily apply to commercial parties.[63] In the context of property guardianship – straddling the boundary between residential occupation in fulfilment of an overarching commercial purpose – the question can be legitimately posed as to how the court negotiates the *Street*-mandated vigilance to fulfil a social purpose, while also seeking to give effect to the bargain from a results-based, purpose-driven perspective. The difficulty is coloured by Lord Templeman's narrow dichotomy of status in the residential context as between tenant or a lodger, dependent upon the provision of attendance and services to denote the latter, but to still connote residential occupation as only falling within these categories and to found tenant status by process of elimination requires a more fulsome consideration. The risk of, at best, a premature (at worst, arbitrary) foreclosure of analysing the purpose is present in Lord Hoffmann's observation that classification does not depend upon an intention additional to that in the agreement itself (the 'relativity point').[64] However, not all residential cases are redolent of the 'landlordism' which requires protecting against.

Roynon and *Khoo* both differ considerably from, for example, *Clarke*, where the 'totality, immediacy, and objectives of the powers exercisable by the council' to temporarily house homeless individuals (and the extensive controls over what went on in the hostel) amplified the legitimacy of not granting the right of exclusive possession.[65] It is a paradigm example of how the nature and furtherance of the statutory

60 Ibid [26] (Arden LJ).

61 Ibid [28] (Arden LJ), [42] (Buxton LJ).

62 M Haley, 'Licences of business premises: contract, context and the reach of Street v Mountford' (2013) 64(4) Northern Ireland Legal Quarterly 425, 426.

63 *Clear Channel UK Ltd v Manchester CC* [2006] 1 EGLR 27, [28]–[29] (Jonathan Parker LJ).

64 *Bruton v London & Quadrant Housing Trust* (n 7 above) 413G.

65 *Westminster City Council v Clarke* (n 19 above) 300H–302C and 302A (Lord Templeman).

duty to house the homeless in hostel accommodation was categorically reflected in the restrictive contractual terms.[66] It is understandable that the courts are reluctant to sanction a licensor's charter in the private sector. However, a results-based, purpose-driven approach does exist in the private sphere as gleaned from all the admissible evidence and draws parallels with the approach in the public sector, and so the categories referred to in earlier cases should not be exhaustive of the type of surrounding circumstances preventing exclusive possession. When thinking about the purpose property guardians serve in filling a real, temporary need to secure property, this should be borne in mind such that the agreed contract was 'real' and did what it set out to do.[67] While not easy to gainsay what a court will decide in the residential context, it may require revisiting the types of exceptional categories identified in *Bruton v London & Quadrant Housing Trust* in the Court of Appeal:[68] that property guardians' occupation may be referable to some other *legal* relationship and can draw upon the extensive jurisprudence determining whose occupation is ministerial. The service occupancy cases show how the primacy of the fulfilment of obligations owed for the discharge of duties *qua* employer/employee will render the occupation a licence.[69] In this way, the agreement and its outworking are attributable to the mutual understanding of the superior purpose, akin to the juridically recognised exceptions of 'other legal relationships' to which property guardianship can be admitted, reflecting how the occupation entails the *absence* of exclusive dominion and control.

Property guardianship entails dwelling in the guarded property as a consequence of the overarching, commercial purpose of providing security for the proprietor.[70] One explanation of property guardians' status can be seen, by analogy, via *Camden LBC v Shortlife Community Housing*[71] where the overarching purpose of providing temporary housing in local authority property stock to a co-operative justified the licensor/licensee conclusion. Underlying the concept was achieving maximum usage of the stock to provide housing in the interim period before redevelopment. Licences rather than tenancies were granted to facilitate a quick turnaround so the property could be handed back to the proprietor, and for occupants to be rehoused elsewhere.

66 W Barr, 'Leases: rethinking possession against vulnerable groups' in E Cooke (ed), *Modern Studies in Property Law* vol IV (Hart 2007) 119, 137.

67 *MacNiven (Her Majesty's Inspector of Taxes) v Westmoreland Investments Limited* [2001] UKHL 6, [2003] 1 AC 311, [40]–[41] (Lord Hoffmann).

68 [1998] QB 834, 843C–D (Millett LJ).

69 *Glasgow Corporation v Johnstone* [1965] AC 609; *Norris v Checksfield* [1991] 1 WLR 1241 cf *Smith v Seghill Overseers* (1874–75) LR 10 QB 422.

70 *Khoo* (n 2 above) [25].

71 (1992) 25 HLR 330.

A true construction of the 'terms of the documents, *the purposes of the transactions, and the surrounding circumstances*' meant the professed intentions were conclusive of the parties' true intentions to not grant exclusive possession of the flats.[72] The label bore no disguise, but spoke to the reality that the council needed to retain possession; the designation of shared communal spaces or reservation of the right to introduce new occupants did not constitute a sham, but ensured that the premises remained under the council's control to fulfil its objectives.[73] To not be cognisant of all the circumstances feeding into the professed intention would rewrite the parties' intentions in favour of some other 'truth' – replacing wholesale the commercial sense of the bargain which would not be the appropriate means to fulfilling the common purpose.[74]

Amplifying the absence of intention to grant exclusive possession in *Camden* was the paramountcy of occupation in relation to rate relief. Rate relief was granted on the basis that the co-operative was in occupation and its charitable status conferred entitlement to relief from the rateable valuation.[75] The co-operative could not rely on its members' individual dwelling in the premises to attract the domestic rate. A similar point can be made with guardianship, and the conclusion drawn in *Camden* has been considered in the Valuation Tribunal for England.[76] In *Ludgate* the guardian company (VPS) accommodated guardians to reside in the property, but that was not the primary purpose: it was the relationship between proprietor and VPS which determined the valuation.[77] The purpose of VPS's licence was that the company would in turn licence guardians to fulfil the provision of security. That the guardians were living in the premises was 'an *additional object* … to achieve that purpose' and therefore did not constitute separate hereditaments.[78] The incongruence between the nature of the building and the legal framework Ludgate House sought to impose meant the guardianship scheme precluded the premises from being occupied for domestic purposes. Admittedly, this will be a relevant factor if the premises guardians occupy is not residential, but nevertheless can be instructive in the court's consideration of the nature, mode and purpose of occupation.

72 Ibid 345 (Millet LJ) (emphasis added).
73 Ibid 345–347.
74 Ibid.
75 Ibid 334. The council's continued claim of housing subsidy meant it could not have taken the properties out of its direct control (at 346).
76 *Ludgate House Ltd v Ricketts & London Borough of Southwark* [2018] 584029623163/537N10 (VTE) reversed *Ludgate House Ltd v London Borough of Southwark* [2019] UKUT 278 (LC).
77 Ibid [40].
78 Ibid (emphasis added).

It is worth again revisiting *Bruton* which looms so large in the property guardianship context. There, a short-life scheme similarly breathed new life into land use. However, despite acknowledging the outer purpose, this was not a special circumstance sufficient enough to justify departing from the *prima facie* agreement of the right of exclusive possession. That the only limitations upon the occupation were for inspections amplified the ongoing paradox that insertion of these terms reflects a need to reserve the limited right of control that, without which, would pass as part of the possessory rights' bundle to the temporary 'owner'. If we draw upon the latest decision in the *Ludgate House* litigation, this time in the Court of Appeal, then *Bruton* may possibly be vanquished owing to the way in which the retention of control operated on the facts.[79] The courts have hitherto been focused on the actual exercise of the retained rights of control envisaged in the agreement and using the absence of such as evidence of sham/pretence, whereas the test should be focused on the effect of such control (where provision is made) *if* exercised and whether it would intervene with the actual occupation of the land concerned (compared with the more superficial control in the contract alighted upon in *Roynon*). It is not unforeseeable that the test of control in *Ludgate* (CA) may incentivise grantors to embark upon moving occupants around and interfering with their rights to the use and enjoyment of the property in order to circumvent the legislation, but, as this article considers, whether such attempts are brazenly motivated by avoidance or whether genuine should turn on context and by looking closely at the nature of the property in question.

The relevant background needs to be viewed against the reality of the guardian scheme *per se* and how 'essential to [its] operation' that (similar to *Camden*) premises be returned swiftly and very well at short notice.[80] The scheme (like *Ludgate House*) serves a commercial purpose, and its continued operation requires giving expression to the agreement taking 'both a textual and contextual approach'.[81] The argument here is not to suggest that all agreements should be found to be licence agreements, but rather that the default starting position should not presume/infer that the agreement was a pretence and cannot be justifiable in every case. While not going back to the bad days of awarding marks for drafting, the words should be given their ordinary effect where there is no ambiguity.[82]

---

79 *London Borough of Southwark v Ludgate House Ltd* [2020] EWCA Civ 1637 (Ludgate (CA)).

80 *Khoo* (n 2 above) [28].

81 Ibid [29]; S Bright, *Landlord and Tenant Law in Context* (Hart 2007) 60.

82 *Malenesian Mission Trust Board v Australian Mutual Provident Society* [1996] UKPC 53, [8] (Lord Hope); Bright et al (n 36 above) 105, 116.

Lord Templeman's reasoning in *Street* has led to interpreting the requirements of a lease in serial fashion rather than as a composite consideration of the true nature of the agreement. Serialisation maintains the spirit of the decision, but in overlooking/diminishing certain other aspects of the reasoning compartmentalises the *indicia* such that the first consideration becomes the fact of exclusive possession and then sham, but this is not the proper test. While sham/pretence empowers the court to declare ineffective certain terms in the agreement at the time the case comes to court, the doctrine requires vigilance, and it should not axiomatically follow that non-operation of a clause is evidence of not intending to be bound at all. While there may be an incongruence between the agreement and what occurs (or is yet to) on the ground, it should not mean the occupier has the *right* to exclude the world at large to give rise to a tenancy. Rather, the starting point should be to consider that the agreement reflects the rights and obligations envisaged by the parties and should not be easily disregarded.[83] Background facts can be instructive in ascertaining whether the right of exclusive possession has been granted, and a results-based approach, revealed by the contract's terms, can offer a way forward for the sake of doctrinal coherence. The integrated-compound approach of subjective–objective intention uses background, agreement and effect as mutually reinforcing each other and can be both influential and (barring sham/pretence) conclusive of status. Formalism and realism can often be viewed as competing paradigms – as though the written agreement is divorced from the reality of what the parties wanted to create. The issue in the property guardianship context is that the realism is not solely that there is residential accommodation (a consequence), the realism is that both parties are aware of the nature and purpose of the occupancy and finds its expression in/given effect by the formal agreement which should take the court away from the cynicism of fork and spade. The *Bruton*-approach in its avowed adherence to *Street* closes off the possibility of considering the nuances of the agreement to over-generalise rather than facilitating the parties' own 'local law' accommodating purpose as part of the regulatory sphere.[84]

Interestingly, Bruton features attempts (albeit slender) by Lords Slynn and Jauncey to engage with the *dicta* of Slade LJ in *Family Housing Association v Jones*[85] where 'misgivings' about the

83 *AG Securities v Vaughan; Antoniades v Villiers* (n 16 above) 475E (Lord Jauncey): '[I]t would not be right to look at the agreements without regard to the circumstances which existed at the time when they were entered into.'

84 J Cartwright, *Contract Law: An Introduction to the English Law of Contract for the Civil Lawyer* 3rd edn (Oxford University Press 2016) 98.

85 [1990] 1 WLR 779.

consequences of the decision would diminish the choice available to providers in the short-life context to best achieve the temporary accommodation purpose.[86] Cognisance of the need to contribute to the efficacy of the parties' purposes should lead the court to give effect to the 'other interest' where the agreement negatives conferring the right of exclusive possession.[87] This 'statement of principle' (as Slade LJ describes it) would be regarded as too pragmatic given that those cases following *Street* and *Bruton* are too far gone to not look beyond the fact that the provision is residential (albeit temporary). The 'expectation gap' within the agreement is (too readily) in-filled with language of 'effect' that functions rather deductively: for the sake of a clearer taxonomy and a 'socially desirable and eminently sensible' jurisprudence,[88] it will be necessary for the courts to avoid continuing to conflate achieving a certain stated, common purpose with those labelling cases overtly attempting to contract out of statute to allow the agreement to stand on its own merits. This would be consistent with *Street* and retains construing the meaning and effect of the agreement to ascertain the proper status of the occupant. The 'purpose-driven' approach would be inductively congruent with *Street*'s articulation of surrounding circumstances/relevant background and would approach the nature and mode of occupancy through a lens of complementarity. It may be too early to tell, but when read together, *Khoo* and *Ludgate* (CA) may illustrate the development of a jurisprudence where the agreement, its effect and nature of occupancy coincide to justify the finding that no exclusive possession is enjoyed by a property guardian given the ministerial, outer purposes of the scheme, and this is both justified and defensible. However, the approach advocated here, while beneficial for doctrinal clarity, would leave property guardians themselves more, not less, vulnerable in the search for the pinnacle of protection provided by legislation.

CODA

Returning to a *laissez-faire* liberalism needs weighing against the social purpose of curtailing freedom of contract in the housing sphere. A pragmatic approach as advocated thus far would be a considerable setback for the protections built up through successive Rent Acts (though itself set back by deregulation in the 1980s). *Khoo* and *Ludgate* (CA) take deregulation even further; meanwhile cases in the aftermath

86 Ibid 793B–D and 793F–G cf *Eastleigh Borough Council v Walsh* [1985] 1 WLR 525. Slade LJ's misgivings had existed for some time: [1985] 49 P & CR 324, 332.

87 *Ogwr Borough Council v Dykes* [1989] 1 WLR 295, 302 (Purchas LJ).

88 After *Bruton v London & Quadrant Housing Trust* [1998] 3 WLR 438, 440B (Sir Brian Neill).

of *Khoo*[89] – with similarities to property guardianship – illustrate the risks of prematurely departing from paternalism. In *Del Rio Sanchez*, a university student entered into a licence agreement (so-called) to occupy a room paying a monthly 'membership fee' (ie rent), and a 'joining fee' for this purported 'accommodation club' (ie the deposit). The judge found the claimant had been granted exclusive possession of the room and that the labels used in the agreement were a sham.[90] In a case such as this, it is easy to see why the court will intervene – this was a labelling case: terms contrived to call a five-pronged instrument a spade and exploit a claimant of considerably less bargaining power. An entirely liberal approach by common law across the rental sector would mark a new-old orthodoxy pre-dating *Street*. Perhaps the way forward lies in the Law Commission's previously recommended 'consumer protection' approach and which, when reflected upon in light of property guardianship, would be timely. The stated aim '*focuses on the contract* between the landlord and the occupier' and approaches the protective regulation in accordance with fairness and transparency rather than depending on the fine technicality between lease/licence status.[91] The simplicity and clarity of reform would bring about (a) the removal of the ambivalence and constructive ambiguity in proving exclusive possession to (b) instead place singular emphasis on the contractual agreement between landlord and occupier for the purposes of fulfilling their statutory obligations.[92] Despite the call for reform, and compared to Wales[93] and Scotland,[94] English law remains the outlier. Given the increasing usage of property guardians, the implications of these reforms are clear: all that would suffice is a contract providing 'evidence of what the parties agreed' – the right to occupy premises as a home.[95] The occupant would be granted either a fixed term or periodic standard contract, and two months' notice would have to be given if seeking to recover possession. For property guardians, the benefits are clear: doing away as a matter of law with the minimal, four-week notice period contained in most guardian agreements. The proposals would capture the reality of the occupation by recognising

89 Eg *Del Rio Sanchez v Simple Properties Management Ltd* (Central London County Court, 24 February 2020).

90 Ibid [59].

91 Law Commission, *Renting Homes* (Law Com No 297, 2006) [1.5] (emphasis added).

92 The centrality of the lease/licence would be retained only insofar as the proprietary character of the parties' relationship is called into question (Law Commission, *Renting Homes 1: Status and Security* (Law Com CP No 162, 2002) [9.39]–[9.40]).

93 Renting Homes (Wales) Act 2016.

94 Private Housing (Tenancies) (Scotland) Act 2016.

95 Law Commission (n 91 above) [3.34].

that (albeit temporary in nature) it is occupied as a home, and not solely for the instrumental purpose of providing security for the proprietor. A consumer approach would at least make a significantly clearer advance for those occupying under a rental agreement by removing the hurdle of having to first prove a landlord and tenant relationship.

CONCLUSION

This article has considered two contrasting cases which offered the first opportunity for the courts to assess property guardians' status in light of the *Street v Mountford* settlement, as impacted by *Bruton v London & Quadrant Housing Trust*. *Khoo* and *Roynon* illustrate how the courts are still working through the finer implications of Lord Templeman's criteria and the perennial tension between pragmatism and principle in giving effect to the parties' agreement, yet in the context of providing residential accommodation, albeit temporarily. Significantly, both judgments illustrate an incomplete analysis of the case law on exclusive possession, and, as a result, the continued fluidity towards factual/legal possession. This article has discussed how a greater degree of nuance can accommodate the parties' complementary purposes and maintain fidelity with the empiricism of English land law which looks at the position on the ground between the parties. In the reality of the *Bruton*- and orthodox tenancy co-existing, taking a compound-integrative approach allows for an appreciation of the purposes underpinning the agreement and a broader, inductive approach to the surrounding circumstances to further support the parties' agreeing to the non-conferral of exclusive possession and justifiable creation of a contractual licence. However, as reform remains elusive and questions concerning property guardianship's place in wider housing law remain outstanding,[96] status still requires to be proved, and guardians can have no property in the very thing they are said to be protecting.

96 *Ludgate* (CA) (n 79 above) [86]–[89] concerning the issue of illegality and property guardianship operating as an unlicensed house in multiple occupation (HMO).

Northern Ireland
Legal Quarterly
Summer Vol. 74.2 (2023) 323–361
Articles DOI: 10.53386/nilq.v74i2.1098

George Gavan Duffy and the legal consequences of the Anglo Irish Treaty, 1921–1923

Thomas Mohr

University College Dublin
Correspondence email: thomas.mohr@ucd.ie

ABSTRACT

George Gavan Duffy (1882–1951) was a signatory of the 1921 'Anglo Irish Treaty'. In the 1930s he enjoyed a notable judicial career and would rise to the position of President of the High Court of Ireland. This article examines a more neglected period of Gavan Duffy's career. It focuses on his brief parliamentary career as a TD in the early 1920s and, in particular, his involvement in the creation of the Constitution of the Irish Free State. This analysis also examines the reasons for the divergence of Gavan Duffy's position from that held by other signatories and supporters of the 1921 Treaty. By late 1922 Gavan Duffy had emerged as a determined critic of the Provisional Government and of the draft Constitution of the Irish Free State that emerged from negotiations in London. This analysis focuses on Gavan Duffy's attempts to amend provisions of the draft Constitution that he believed went further than the strict legal demands of the 1921 Treaty. The conclusion assesses Gavan Duffy's attitude towards the legal consequences of the 1921 Treaty and his attempts to mitigate their impact on the 1922 Constitution of the Irish Free State.

Keywords: George Gavan Duffy; Anglo Irish Treaty; Irish Free State; 1922 Constitution; Constituent Assembly; King; oath; Privy Council; popular sovereignty.

INTRODUCTION

The centenary of the 1921 Anglo Irish Treaty has seen renewed attention to those who negotiated and signed this historic document. These include George Gavan Duffy who would later become a judge of the Irish High Court in 1936 and its President in 1946. Gavan Duffy's role in the negotiation and signing of the 1921 Treaty has been covered by numerous works.[1] His legal career as a barrister in the 1920s and 1930s and his subsequent career as a judge have received significant attention in publications that focus on his advisory role in the drafting

1 For example, see Frank Pakenham, *Peace by Ordeal* (Sidgwick & Jackson 1935); Frank Gallagher, *The Anglo Irish Treaty* (Hutchinson 1965); and Joseph M Curran, *The Birth of the Irish Free State 1921–1923* (University of Alabama Press 1980).

of the 1937 Constitution.[2] Others focus on his legal career that includes involvement in many historic court cases including *State (Ryan) v Lennon*,[3] as a barrister, and *State (Burke) v Lennon*[4] and *Buckley v Attorney General*, as a judge of the High Court.[5] This article examines a comparatively neglected period of his life that fell between the signing of the 1921 Treaty and his subsequent legal career. It focuses on Gavan Duffy's brief parliamentary career as a Teachta Dála (TD) in the early 1920s and, in particular, on his involvement in the creation of the Constitution of the Irish Free State. This aspect of Gavan Duffy's life is of interest because it witnessed a parting of ways with other signatories and supporters of the 1921 Treaty. It also reveals Gavan Duffy's response to the legal consequences of the document that he signed and his attempts to mitigate their impact on the 1922 Constitution of the Irish Free State.[6]

Gavan Duffy sat in the third Dáil Éireann, sitting as a special Constituent Assembly, that would approve the final text of the Constitution. He proved to be the most prolific representative in moving amendments supported by elaborate legal argument. Most of these amendments concerned provisions relating to the 1921 Treaty, and Gavan Duffy made the most of his position as a signatory in promoting them. His fundamental argument was that the draft Constitution had not made full use of the autonomy provided by the 1921 Treaty. This resulted in a serious clash with former colleagues in the Provisional Government and may also have contributed to the truncation of his parliamentary career. This analysis also examines Gavan Duffy's hopes

2 For example, see Dermot Keogh and Andrew McCarthy, *The Making of the Irish Constitution 1937* (Mercier 2007); Gerard Hogan, *The Origins of the Irish Constitution, 1928–1941* (Royal Irish Academy 2012); and Donal K Coffey, *Constitutionalism in Ireland, 1932–1938 – National, Commonwealth and International Perspectives* (Palgrave Macmillan 2018).

3 [1935] IR 170.

4 [1940] IR 136.

5 [1950] IR 67. See also G M Golding, *George Gavan Duffy, 1882–1951 – A Legal Biography* (Irish Academic Press 1982); Frank Connolly, 'An interim assessment of the late Mr Justice George Gavan Duffy as advocate and judge' [1976] 70 Incorporated Law Society of Ireland Gazette 129–130 and 134; Colum Gavan Duffy, 'George Gavan Duffy' (1983) 36(3) Dublin Historical Record 90 reprinted in (2002) 2(2) Judicial Studies Institute Journal 1; and Gerard Hogan, 'George Gavan Duffy' in James McGuire and James Quinn (eds), *Dictionary of Irish Biography* vol 3 (Cambridge University Press 2009) 510–512.

6 Unless stated otherwise, the numbering of all provisions of the 1922 Constitution of the Irish Free State used by this article follows the numbering used when the Constitution officially came into force on 6 December 1922. The numbering of these provisions changed considerably over the various stages of drafting the Constitution and as a consequence of amendments made during the proceedings of the Dáil sitting as a Constituent Assembly.

on how the infant Irish Free State might augment its autonomy in the years that followed. All these events were of lasting significance to Gavan Duffy himself. This reality was reflected in his decision to create a recording of his memories on this subject, in preference to other aspects of his career, shortly before his death in 1951.[7]

GEORGE GAVAN DUFFY (1882–1951)

George Gavan Duffy was the son of Charles Gavan Duffy, a Young Irelander who would emigrate to Australia where he would rise to the position of premier of Victoria. He was one of a large number of gifted children, including his elder half-brother Frank Gavan Duffy, who would become Chief Justice of the High Court of Australia.[8] George was born in England but spent considerable portions of his youth in different European countries including Ireland. He would later move into the legal professions when he qualified as a solicitor in London in 1907. He came to prominence in Irish nationalist circles in 1916 when he assisted in the defence of Roger Casement for his involvement with the Easter rising. Casement would be executed for high treason and Gavan Duffy's role in his defence would result in him being asked to leave his firm of solicitors.[9] The trial deepened Gavan Duffy's involvement in Irish politics and changed the direction of his legal career. In 1917 he moved to Ireland, qualified as a barrister and became increasingly involved with the Sinn Féin party.

Gavan Duffy's parliamentary career officially began when he won a parliamentary seat in Dublin in the 1918 general election. As a member of Sinn Féin he declined to take his seat at Westminster and instead attended the opening of the first Dáil Éireann on 21 January 1919. His linguistic skills ensured that he was chosen to read out the French text of the declaration of independence issued on that day. Gavan Duffy's legal background also ensured that he was appointed to a committee

7 'The Hon Mr Justice George Gavan Duffy' (Military Archives, Voice Recordings).

8 Patrick O'Callaghan, 'Brothers at law: Chief Justice Frank Gavan Duffy and George Gavan Duffy' (2004) 78 Australian Law Journal 738.

9 C Gavan Duffy (n 5 above) 92.

charged with drafting a constitution for the Dáil.[10] However, he had little involvement in Dáil proceedings after the opening sessions as he spent much of the next two years in continental Europe in an unsuccessful attempt to secure recognition for the self-declared Irish republic. This ensured that Gavan Duffy missed the darkest days of the conflict that engulfed Ireland between 1919 and the truce of 11 July 1921.

Gavan Duffy's parliamentary career really began in August 1921 when, having returned to Ireland, he attended the debates of the second Dáil Éireann which was now able to meet unmolested thanks to a truce with Crown forces that began on 11 July 1921. For the first time he made a number of significant parliamentary contributions. He reported to the Dáil on his activities in promoting the *Irish Bulletin*, a newspaper dedicated to supporting the struggle for an Irish republic. Gavan Duffy's speech on this subject made some unfortunate references to difficulties in propagating information in the face of the 'octopus' of news agencies run by 'big Jew firms in London'.[11] His contributions to the debates of the second Dáil also included an unsuccessful protest against Eamon de Valera's appointment as President in addition to his existing role as 'Príomh Aire', or Prime Minister, on the grounds that it was dangerous precedent to allow such a concentration of power.[12] Gavan Duffy's legal background may also have influenced his opposition to aspects of emergency legislation passed by the Dáil.[13]

In October 1921 de Valera accepted an invitation to send delegates to London to negotiate a permanent peace. De Valera made the controversial decision not to join this delegation himself and Arthur Griffith was appointed in his place to lead the Irish delegation. Griffith

---

10 Brian Farrell, 'A note on the Dáil Constitution, 1919' (1969) 4 Irish Jurist 127. The nature of this document has long been disputed with many arguing that it was only intended to be a constitution for the Dáil and not a constitution for Ireland. For example, Seán McBride wrote, in a submission to the New Ireland Forum: 'In addition, the first Dáil adopted "The Democratic Programme of Dáil Éireann", and a "provisional Constitution of Dáil Éireann". ... Neither of these instruments purported to be a Constitution for the Republic.' University College Cork Archives, O'Rahilly papers, U. 118, Box 6, Submission to the New Ireland Forum, 1984. See also Basil Chubb, *The Constitution of Ireland* (Institute of Public Administration 1966) 8 and *The Government and Politics of Ireland* (Oxford University Press 1974) 62–63.

11 *Dáil Debates*, vol S, no 5, col 53, 23 August 1921. John Kelly comments on 'the dark side of a remarkable judge' in assessing Gavan Duffy's judgment in *Schlegel v Corcoran and Gross* (1942) IR 19. J M Kelly, *The Irish Constitution* 2nd edn (Jurist Publishing 1984) 665. See also Ruth Cannon, 'The bigoted landlord: a re-examination of *Schlegel v Corcoran and Gross*' (2005) 27 Dublin University Law Journal 248; and O'Callaghan (n 8 above) 741–742.

12 *Dáil Debates*, vol S, no 5, col 55, 23 August 1921.

13 For example, see *Dáil Debates*, vol 3, col 2526–2532, 26 June 1923.

was joined by two other members of the Dáil Cabinet, Michael Collins and Robert Barton. George Gavan Duffy and Eamonn Duggan were also appointed to the delegation, largely on the basis of their legal expertise.[14] Robert Erskine Childers was appointed as secretary to the Irish delegation. Additional legal support was provided by John Chartres, a barrister, diplomat and former civil servant, and an advisory committee in Dublin that included James Nolan-Whelan, a leading barrister, James Murnaghan, a law lecturer at University College Dublin and future judge of the Irish Supreme Court, and John O'Byrne, a barrister and future attorney general and judge of the Supreme Court.[15]

Frank Pakenham concludes that Gavan Duffy, along with the rest of the Irish legal experts, was outclassed by the formidable Lord Birkenhead, who held the office of Lord Chancellor in Lloyd George's Government.[16] In reality, the British offer of Dominion status demanded a knowledge of British imperial law and the constitutional law of Dominions, such as Canada, that few Irish lawyers possessed. Birkenhead had far greater legal supports available in these areas. Erskine Childers had a far better grasp of this field than Gavan Duffy or Duggan thanks to his authorship of his 1911 work *The Framework of Home Rule*, which considered the possibility of granting Dominion status to Ireland in some detail.[17] However, Childers did not have the official status of a delegate and, in any case, by 1921 he was now firmly committed to seeking a political settlement that went far beyond Dominion status.[18] Gavan Duffy along with other members of the delegation attempted to close their gap in knowledge by studying works dedicated to the legal implications of Dominion status.[19] Nevertheless, Gavan Duffy's knowledge of this area of law was always hampered by an intense hostility to a Dominion settlement for Ireland which may have resulted in a lack of objectivity in assessing its legal consequences.

Gavan Duffy embraced de Valera's scheme for 'external association' for Ireland, a compromise that offered far greater autonomy than the

14 Frank Pakenham, *Peace by Ordeal* (Sidgwick & Jackson 1972) 84.

15 *Dáil Debates*, vol T, no 4, col 201, 16 December 1921.

16 Pakenham (n 14 above) 106.

17 (Edward Arnold, 1911).

18 Childers admitted that he had 'passed through the Dominion phase' but concluded that going back to it would be 'an almost impossible and unthinkable thing'. *Donegal Democrat* (Ballyshannon, 1 December 1922) 2.

19 This would focus on H Duncan Hall, *The British Commonwealth of Nations – A Study of its Past and Future Development* (Methuen 1920). See Pakenham (n 14 above) 226; and H Duncan Hall, *Commonwealth – A History of the British Commonwealth of Nations* (Van Nostrand Reinhold 1971) 198 and 912.

familiar model of Dominion status.[20] He persisted in championing this constitutional compromise despite repeated British rejections of it during the negotiations in London. Gavan Duffy remained convinced right up to the fateful moment of signing the Treaty that the British delegation would, in the end, accept greater autonomy for Ireland than their repeated offer of Dominion status. He was deeply concerned at the emergence of sub-conferences during the Treaty negotiations which tended to exclude him and other delegates from important negotiations led by Griffith and Collins.[21] His protests proved to be in vain, a reflection of his lesser standing within the Irish delegation.

Gavan Duffy's subordinate status did not prevent him from decisively influencing the final days of the negotiations. He joined Barton in insisting on a final push for external association in a meeting with the British delegation on 4 December 1921. Michael Collins refused to join this initiative and Arthur Griffith reluctantly agreed to accompany Gavan Duffy and Barton at the last minute. Once again, the British delegation rejected an Irish offer based on external association. Gavan Duffy continued to press this option in the face of British insistence on Dominion status and concluded: 'We should be as closely associated with you in all large matters as the Dominions, and in the matter of defence still more so; but our difficulty is coming into the Empire.'[22] This blunt refusal to remain within the British Empire had an electric effect on the British delegation and undermined Irish negotiating strategy. Lloyd George's negotiating stance in the discussions that followed included elements of theatrics, but the reality that the talks were reaching their limits had become all too clear.

On the evening of 5 December 1921 three of the five Irish delegates, Griffith, Collins and Duggan, decided to sign the draft Treaty. The hours that followed, once again, highlighted the secondary nature of Gavan Duffy's membership of the delegation. The three delegates focused their attention on Barton in their efforts at persuasion while largely ignoring Gavan Duffy. Barton was, after all, a member of the Dáil Cabinet, and it was assumed that if he gave way Gavan Duffy would be forced to follow. This assumption proved all too accurate and Gavan Duffy agreed to sign soon after Barton gave in to persuasion.

Gavan Duffy was placed in an unenviable position of having to publicly defend a settlement that he had signed with such reluctance in the Dáil debates that followed. He openly admitted to the Dáil: 'I am not enthusiastic about this Treaty although I am going to support

20 See Thomas Mohr, 'The Anglo Irish Treaty – legal interpretation, 1921–1925' (2021) 66 Irish Jurist 1.

21 Pakenham (n 14 above) 145, 190 and 202.

22 Ibid 217.

it.'[23] On another occasion he declared: 'I am going to recommend this Treaty to you very reluctantly, but very sincerely, because I see no alternative.'[24] He did admit that the Treaty offered gains to the Irish people, and, in particular, stressed that 'this Treaty gives them what they have not had for hundreds of years; it gives them power, it puts power of control, power of Government, military power in the hands of our people and our Government'.[25] Yet, Gavan Duffy also devoted a considerable portion of his speeches to the perceived shortcomings of the Treaty. He placed particular emphasis on acceptance of the King which he insisted had inflicted a 'grievous wound upon the dignity of this nation'.[26] Gavan Duffy also devoted a considerable portion of his speech to outlining the circumstances under which his signature of the Treaty had been 'extorted' from him.[27] He joined Robert Barton in arguing that Lloyd George's ultimatum, stressing the possibility of renewed conflict, ensured that their signatures had been made under duress. Yet, he added that duress was not sufficient cause to reject the Treaty and insisted that its opponents produce a 'rational alternative'.[28] His appeal to opponents of the settlement was also evident in his conclusion: 'My heart is with those who are against the Treaty, but my reason is against them, because I can see no rational alternative.'[29] Gavan Duffy's position was sincere but opponents of the proposed settlement lost no time in making capital from it. Mary MacSwiney summarised his stance on the Treaty by concluding: 'He has given weak support to it, but he has acknowledged it is a very pitiful instrument indeed, but that it is better than war.'[30]

CABINET MINISTER

The Dáil finally approved the Treaty settlement by 64 to 57 votes. This opened the way towards the formation of an Irish Provisional Government that would oversee the handover of the 26 counties of the south and west of Ireland until the official formation of the Irish Free State on 6 December 1922. Michael Collins led the Provisional Government that was recognised by the British Government while Arthur Griffith led a parallel Dáil Cabinet that could be recognised by Irish opponents of the Treaty.

---

23 *Dáil Debates*, vol T, no 4, col 184, 16 December 1921.
24 *Dáil Debates*, vol T, no 8, col 85, 21 December 1921.
25 Ibid.
26 Ibid.
27 Ibid.
28 *Dáil Debates*, vol T, no 8, col 87–888, 21 December 1921.
29 Ibid.
30 *Dáil Debates*, vol T, no 8, col 114, 21 December 1921.

Robert Barton had now moved closer to opponents of the Treaty and consequently could not be considered for a ministerial position in either the Provisional Government or the Dáil Cabinet. Eamonn Duggan was appointed as Minister for Home Affairs, possibly on the basis of his legal background. George Gavan Duffy was appointed as Minister for Foreign Affairs, a promotion that was probably inspired by his linguistic skills and his experience in trying to gain recognition for the self-declared Irish republic between 1919 and 1921. However, his status as a signatory of the Treaty may also have been an important consideration behind this appointment. The signatories of the Treaty had already lost Barton and did not want to lose another of their number.

As Minister for Foreign Affairs, Gavan Duffy decided to join the Dáil Cabinet but not the Provisional Government. Gavan Duffy justified this position on the basis that it was best to keep his embryonic Department of Foreign Affairs autonomous in the hope that it could continue to function should relations break down between London and the Irish Provisional Government.[31] This decision may also have reflected Gavan Duffy's personal preferences and the reality that the nascent Department of Foreign Affairs had little role to play in the transition of power overseen by the Provisional Government. Gavan Duffy would later emphasise that he had never been part of the Provisional Government to disclaim all responsibility for the draft Irish Constitution that was negotiated with London.[32]

One of the most noteworthy features of Gavan Duffy's brief career as a cabinet minister was his obvious discomfort with this role. He made repeated threats of resignation over a period of just over six months. His first threat of resignation occurred in February 1922, less than a month after his appointment, over his offence at a newspaper article written by Collins.[33] As will be seen at a later stage, he would make a second threat to resign in June 1922 and would finally follow through on a third threat to resign in July 1922.

DRAFTING THE CONSTITUTION OF THE IRISH FREE STATE

The 1921 Treaty provided little guidance on the means by which the Constitution of the Irish Free State would be brought into being. The British Government indicated soon after the signing of the 1921 Treaty that the Irish Provisional Government could draft its own Constitution

31 *Dáil Debates*, vol 1, col 531, 21 September 1922.

32 Ibid.

33 University College Dublin (henceforth UCD) Archives, Gavan Duffy Papers, P152/196, Gavan Duffy to Collins, 6 February 1922.

as long as it was compatible with the terms of the Treaty and certain promises made to the southern Protestant minority.[34] These conditions rendered it necessary to provide the British Government with a confidential preview before the draft Constitution was made public. This reality was conceded by the Provisional Government in February 1922 when the Constitution was still in the early stages of drafting.[35]

The Provisional Government appointed a special Constitution Committee to create the preliminary drafts of the future Constitution. It produced three alternative draft Constitutions known as Drafts A, B and C. The Provisional Government chose Draft B, which formed the basis of the draft Constitution and reviewed and amended it until the end of May 1922. The draft settled by the Provisional Government and its advisors, known as Draft D, was later brought to London for the anticipated review.

Gavan Duffy played a secondary role in the drafting of the Constitution notwithstanding his legal experience. He was not chosen to serve on the Constitution Committee, and he did not form part of the delegation that brought Draft D to London for the difficult negotiations that followed.[36] Nevertheless, claims that Gavan Duffy was not consulted during the drafting of the Constitution are inaccurate.[37] He was shown early drafts and consulted by members of the Constitution Committee in preference to other cabinet ministers.[38] Gavan Duffy, in turn, attempted to render practical assistance to the Constitution Committee, for example in recommending the consultation of legal experts from the Dominions.[39] He also had official access to the evolving text of the draft Constitution when the Constitution Committee finally sent drafts for consideration by cabinet ministers. He was also kept up to date on the negotiations with the British Government in May and June 1922.

Gavan Duffy appears to have shared the general approach taken by the Provisional Government in seeking as short and simple a draft

34 This was confirmed in a letter sent by Lloyd George to Griffith on 13 December 1921 that was later read out in the Dáil. *Dáil Debates*, vol T, no 6, col 21–22, 19 December 1921.

35 National Archives of Ireland (henceforth NAI), Cabinet Minutes, G1/1 2 February 1922 and The National Archives (henceforth TNA), CAB 43/6 22/N/60(6), meeting between the British and Irish signatories, approval of draft Constitution, 26 February 1922.

36 Golding expresses some surprise that Gavan Duffy was not asked to join the Constitution Committee (n 5 above) 24.

37 Mary Kotsonouris, 'The George Gavan Duffy Papers' (2000) 8(4) History Ireland.

38 For example, see National Library of Ireland (henceforth NLI), James Green Douglas Papers, Ms 49,581/8/1, George Gavan Duffy to James Douglas, 27 February 1922.

39 NLI, James Green Douglas Papers, Ms 49,581/7/12, George Gavan Duffy to James Douglas, 26 January 1922.

Constitution as possible. He produced a detailed analysis of Draft B with the conclusion that it was too long and recommended substantial changes. Gavan Duffy recommended that 27 of the 81 articles contained in Draft B be cut entirely and recommended amendments to all but one of the remaining articles.[40] Although Gavan Duffy would later reveal himself as a supporter of the use of religious and natural law principles in the interpretation of constitutional law, there is no evidence that he, as a cabinet minister, pushed for the inclusion of such principles in the 1922 Constitution.[41]

The impact of the 1921 Treaty on the draft Constitution depended on the interpretation of article 2 of the former which linked the constitutional status of the Irish Free State to that of Canada in terms of 'law, practice and constitutional usage'. In early 1922 Irish cabinet ministers and their advisors interpreted the provisions of the Treaty as allowing them to unite Dominion law and practice in order to create constitutional provisions that were entirely different from those of the existing Dominions. They believed that a synthesis of 'law' and 'practice' would allow the creation of an Irish Constitution that reflected constitutional realities in the existing Dominions and so avoid inclusion of obsolete legal provisions included in their Constitutions.[42] However, Collins and Griffith went further in instructing the Constitution Committee to ignore legal institutions reflected in the Constitutions of the existing Dominions that were not actually obsolete such as the appeal to the Judicial Committee of the Privy Council.[43]

The Constitution Committee followed these instructions and produced draft Constitutions that largely ignored the Treaty settlement and Dominion precedents.[44] The apparent objectives were to achieve as much autonomy as possible for the future Irish Free State and to produce a Constitution that opponents of the Treaty might be able to accept. Once again, Gavan Duffy was in entire agreement with this approach. In March 1922 he wrote to Collins to argue that if the

40 UCD Archives, Gavan Duffy Papers, P152/199, Gavan Duffy to Collins, 6 April 1922.

41 On Gavan Duffy's support for natural law, see Thomas Mohr, 'Natural law in early twentieth century Ireland – State (Ryan) v Lennon and its aftermath' (2021) 42(1) Journal of Legal History 1.

42 For example, see *Dáil Debates*, vol T, no 6, col 47, 19 December 1921. See also Brian Farrell, 'The drafting of the Irish Free State Constitution' (1970) 5 Irish Jurist 115, 343 and (1971) 6 Irish Jurist 111, 345; Curran (n 1 above) 200–218; and D H Akenson and J F Fallin, 'The Irish Civil War and the drafting of the Free State Constitution' (1970) 5(1) Eire–Ireland 10; 5(2), Eire–Ireland 42 and 5(4) Eire–Ireland 28.

43 NAI, Department of the Taoiseach, S8952, Constitution Committee, report of first meeting, 24 January 1922.

44 Farrell (n 42 above) 345; Curran (n 1 above) 200–218; and Akenson and Fallin (n 42 above).

draft Constitution was 'as good as the Treaty ... allows it to be, it will knock the bottom out of the Opposition and ... should give a priceless opportunity of uniting the country'.[45]

The decision to minimise the impact of the 1921 Treaty and Dominion precedents ensured that Draft D omitted provisions reflected in the Constitutions of other Dominions, including the parliamentary oath, certain functions of the King and his representative, powers to delay and veto legislation and the appeal to the Judicial Committee of the Privy Council. The text made little more than token references to the Treaty, the representative of the Crown and relations with the Commonwealth in a short section dealing with 'External Affairs'.[46] This approach ensured that the British–Irish negotiations on the draft Constitution, that would take place in May and June 1922, would be a sequel, or even a continuation, of the negotiations that preceded the signing of the Treaty. Gavan Duffy's contributions to the Dáil debates on acceptance of the Treaty in late 1921 illustrate that he was already anticipating the possibility of a future Constitution reversing some of the less acceptable aspects of the settlement that he had just signed. He told the Dáil that the drafters of the Constitution could and should 'relegate the King of England to the exterior darkness as far as they can'.[47]

One of Gavan Duffy's most interesting recommendations on the draft Constitution concerned his desire to exclude the only provision of Draft B that made any reference to the Treaty settlement. He recommended the removal of a provision within article 78 of Draft B that recognised that the 1921 Treaty would enjoy force of law. Gavan Duffy argued that this was 'unnecessary and very undesirable'.[48] The only concession he was prepared to make to the Treaty settlement was a recognition that the representative of the Crown should sign Irish legislation while making it clear that this person would have no power to refuse to do so. Even here, Gavan Duffy was not prepared to allow this into the Constitution proper and recommended its inclusion in a rider to the Constitution.[49] He was broadly in agreement with Collins and Griffith in desiring to minimise the impact of the 1921 Treaty on the text of the draft Constitution but went further in putting this approach into practice.[50]

---

45 UCD Archives, Gavan Duffy Papers, P152/197, Gavan Duffy to Collins, 10 March 1922.

46 NAI, Department of the Taoiseach, S8953, article 78 of Draft B.

47 *Dáil Debates*, vol T, no 8, col 86, 21 December 1921.

48 NAI, Constitution Committee, V13, suggested amendments to the proposed Constitution, 11 April 1922.

49 Ibid.

50 NAI, Department of the Taoiseach (n 43 above).

BRITISH–IRISH NEGOTIATIONS ON THE DRAFT CONSTITUTION

Gavan Duffy had always argued against giving the British Government a confidential preview of Draft D before it was released to the public.[51] However, this course would have ensured that inevitable British dissatisfaction with Draft D, which contained no more than token concessions to the Treaty settlement, would have been played out in public rather than in private. The results of the difficult bilateral negotiations over the draft Constitution, assuming that they occurred at all, would also have been exposed to the full glare of public scrutiny. It is unlikely that the Treaty settlement would have survived such an ordeal.

The final text of Draft D was taken to London on 27 May 1922. Gavan Duffy was not invited to join the Irish delegation that travelled to London, which included Arthur Griffith, Michael Collins and Hugh Kennedy as legal advisor. Gavan Duffy was concerned that the Irish delegation, apart from Kennedy, appeared not to fully appreciate the challenge facing them with respect to the draft Constitution. He anticipated that the British Government would be 'horribly frightened by the bad example that the Constitution gives to the independent elements of the Dominions'. He warned that the British Government would make 'desperate efforts' to insert provisions derived from the 1921 Treaty, in particular the parliamentary oath, into the draft Constitution but recommended that Irish negotiators maintain an 'unyielding attitude' in the face of 'idiotic British sentiment'.[52]

The British Government was quick to reject the draft Constitution presented to them by the Provisional Government on the basis that it was incompatible with the demands of the 1921 Treaty. Lloyd George despaired that they had gone back to the very first day of the Treaty negotiations. He concluded that the draft Irish Constitution was 'a complete evasion of the Treaty and a setting up of a Republic with a thin veneer'.[53] The result was an immediate rise in tensions. The Irish delegation was told that the draft Constitution reflected a refusal to accept Dominion status and was a direct negation of the Treaty signed just over six months earlier.[54]

---

51 UCD Archives, Gavan Duffy Papers, P152/197, 10 March 1922. See Military Archives (n 7 above).

52 UCD Archives, Gavan Duffy Papers, P152/202, Gavan Duffy to Collins, 26 May 1922.

53 TNA, CAB 43/1 22/N/148(3), meeting of British signatories, 27 May 1922 and CAB 43/7 22/N/162, twenty-fourth meeting of the British signatories (S.F.B.) 24th Conclusions, 27 May 1922.

54 UCD Archives, Gavan Duffy Papers, P152/205, minutes of Cabinet meeting, 3 June 1922.

The British Government began to consider the possibility of a collapse of the Treaty settlement and began to explore possible military responses.[55]

Collins returned to Dublin in early June and proposed to Irish cabinet ministers that they sidestep British objections by publishing a 'skeleton' Irish Constitution with the text of the 1921 Treaty attached in a schedule.[56] Gavan Duffy proved to be a strong supporter of this scheme and insisted that even the token references to the Treaty and Commonwealth that appeared in Draft D might be omitted from a short Constitution.[57] Nevertheless, this scheme was soon abandoned as impractical. It anticipated further British–Irish negotiations to fill in details left unsettled by the skeleton Constitution. These piecemeal negotiations on the final form of the Constitution would have proceeded under the full glare of publicity which, once again, would have been deeply embarrassing for both governments.[58]

Tensions were reduced when both sides agreed to a redrafting of Draft D. This meant enshrining a number of key provisions into the Irish Constitution that mirrored those of the existing Dominions while following Dominion practice in the interpretation and application of these provisions.[59] The result was the inclusion of specific constitutional provisions dealing with the position of the King, the representative of the Crown, the parliamentary oath and the appeal to the Judicial Committee of the Privy Council. Gavan Duffy objected to acceptance of these 'offensive shibboleths'[60] and concluded that they 'concede more than we are compelled to concede to England by our obligations under the Treaty'.[61] He was particularly offended by the addition of

55 For example, see TNA, CAB 43/1 22/N/148(1), conference of British representatives, 23 May 1922 and CAB 27/153, PGI 21st Conclusions.

56 UCD Archives, Gavan Duffy Papers, P152/204, minutes of Cabinet meeting, 2 June 1922.

57 UCD Archives, Gavan Duffy Papers, P152/208, Gavan Duffy to Collins, 7 June 1922.

58 Gavan Duffy advocated placing the following in a preamble to the skeleton Constitution: 'Some differences of opinion as to the character of those provisions having developed between the British signatories of the Treaty and the Irish representatives, the Provisional Government is making every effort to arrive at a fair and harmonious solution, and the proposed supplementary provisions will be published as soon as possible.' UCD Archives, Gavan Duffy Papers, P152/208, Gavan Duffy to Collins, 7 June 1922.

59 For example, see articles 41, 51 and 60 of the 1922 Constitution of the Irish Free State.

60 *Irish Independent* (Dublin, 25 September 1922) 6.

61 UCD Archives, Gavan Duffy Papers, P152/210, Gavan Duffy to Griffith, 19 June 1922.

a provision that would be known as the 'Repugnancy Clause',[62] that made clear that any aspect of Irish law that was incompatible with the terms of the 1921 Treaty would be rendered 'absolutely void and inoperative'.[63] Gavan Duffy concluded that this provision was 'thoroughly rotten', made the Treaty part of the Constitution and also ensured that the Constitution derived its force from the Treaty.[64]

The amended version of the draft Constitution was finally published in the newspapers on 16 June 1922. Three days later Gavan Duffy wrote to Arthur Griffith declaring that he could no longer support the draft Constitution and made clear his intention to resign as Minister for Foreign Affairs.[65] Griffith managed to persuade him to delay his resignation until the opening of the Constituent Assembly.[66] In the interim, Gavan Duffy was told not to discuss the draft Constitution until the Constituent Assembly began its work.[67] Finally, the Government's decision to commence the winding-up of the Dáil courts proved too much for Gavan Duffy who brought forward his resignation to July 1922.[68]

---

62 UCD Archives, Gavan Duffy Papers, P152/210, Gavan Duffy to Collins, 14 June 1922. See s 2 of the Constitution of the Irish Free State (Saorstát Éireann) Act 1922 (Dublin) and preamble to the Irish Free State Constitution Act 1922 (Session 2) (Westminster). The use of the term 'Repugnancy Clause' for these parallel provisions was created by Leo Kohn. See Leo Kohn, *Constitution of the Irish Free State* (Allen & Unwin 1932) 98. See also Article 50 of the 1922 Constitution of the Irish Free State.

63 The Repugnancy Clause provided: 'The said Constitution shall be construed with reference to the Articles of Agreement for a Treaty between Great Britain and Ireland set forth in the Second Schedule hereto annexed (hereinafter referred to as "the Scheduled Treaty") which are hereby given the force of law, and if any provision of the said Constitution or of any amendment thereof or of any law made thereunder is in any respect repugnant to any of the provisions of the Scheduled Treaty, it shall to the extent only of such repugnancy, be absolutely void and inoperative and the Parliament and the Executive Council of the Irish Free State (Saorstát Éireann) shall respectively pass such further legislation and do all such other things as may be necessary to implement the Scheduled Treaty.' Constitution of the Irish Free State (Saorstát Éireann) Act 1922 (Ir), s 2 and preamble to the Irish Free State Constitution Act 1922 (Session 2) (UK). See also article 50 of the Constitution of the Irish Free State. The term 'Repugnancy Clause' was introduced by Kohn in *The Constitution of the Irish Free State* (n 62 above) 98.

64 UCD Archives, Gavan Duffy Papers, P152/209, Gavan Duffy to Collins, 14 June 1922.

65 UCD Archives, Gavan Duffy Papers, P152/210, Gavan Duffy to Griffith, 19 June 1922.

66 UCD Archives, Gavan Duffy Papers, P152/211, Griffith to Gavan Duffy, 20 June 1922.

67 Ibid.

68 Golding (n 5 above) 26.

GAVAN DUFFY IN THE CONSTITUENT ASSEMBLY

Gavan Duffy's objections to the draft Constitution continued to preoccupy him after he resigned from the Dáil Cabinet and sat as an independent TD. He was one of the most active members of the Dáil, sitting as a Constituent Assembly, and prepared a substantial number of amendments for the draft Constitution. However, it is worth mentioning that Gavan Duffy touched on many other subjects during his brief parliamentary career, many of which reflected his interests in legal matters. For example, he did not shy away from questioning the legal basis of military tribunals established during the Civil War and was very active in the debates on the emergency legislation that attempted to fill this legal vacuum.[69] He also raised objections during the winding-up of the Dáil courts.[70] Gavan Duffy was deeply concerned with the rights of anti-Treaty prisoners captured during the Civil War and offended the Provisional Government by referring to them as 'political prisoners'.[71] He was firmly opposed to the execution of such prisoners which would become a feature of the conflict. It is difficult to dispute his claim that the execution of four prisoners in Dublin's Mountjoy Prison on 7 December 1922 lacked any legal basis and violated a Constitution that had come into force less than two days earlier.[72] Gavan Duffy's reaction to the execution of Erskine Childers provoked a particularly emotional debate in the Dáil that culminated in W T Cosgrave accusing him of lacking moral courage.[73] The full extent in the collapse of relations became evident in April 1923 when Gavan Duffy's house was raided by the Criminal Investigation Department who were searching for a wanted man, possibly Eamon de Valera. Gavan Duffy's protests in the Dáil met with little sympathy from Kevin O'Higgins who accused him of 'histrionics and mock heroics'.[74] O'Higgins also mocked Gavan Duffy's absence from home at the time of the search by declaring that he had 'gone to his spiritual home abroad'.[75]

The position of the nascent Irish Government had changed beyond recognition when the Constituent Assembly convened in September 1922. A bitter Civil War had erupted on 28 June 1922 that ended all hopes of compromise and reconciliation with opponents of the Treaty. Griffith and Collins had died the following August leaving W T Cosgrave

69 For example, see *Dáil Debates*, vol 3, col 2526–2532, 26 June 1923.
70 *Dáil Debates*, vol 4, col 1310–1321, 24 July 1923.
71 *Dáil Debates*, vol 1, col 206–209, 13 September 1922.
72 *Dáil Debates*, vol 2, col 51, 8 December 1922.
73 *Dáil Debates*, vol 1, col 2363, 28 November 1922.
74 *Dáil Debates*, vol 3, col 591, 3 May 1923.
75 *Dáil Debates*, vol 3, col 592, 3 May 1923.

to take over leadership of both the Provisional Government and the Dáil Cabinet. The removal of any incentive to placate opponents of the Treaty led Cosgrave to unite these two institutions in September 1922. Soon afterwards, he decided to convoke a new Dáil, sitting as a Constituent Assembly, that would amend and formally enact the draft Constitution that had returned from London. The Provisional Government indicated that it would resign unless certain key provisions were enacted without amendment. These provisions, which might be called the 'Treaty articles', included all the provisions concerning the Treaty settlement that had been agreed with the British Government in June 1922.[76] The outbreak of Civil War prevented the attendance of any opponents of the Treaty in the Constituent Assembly. Nevertheless, there were a few TDs who were prepared to challenge the Treaty articles within the draft Constitution. These included a number of independents and some members of the Labour Party. However, the most active and abrasive critic of these Treaty articles within the Constituent Assembly was George Gavan Duffy.

Gavan Duffy was not content to simply criticise the inclusion of the Treaty articles in the draft Constitution. He produced amendments, with accompanying legal justifications, in respect of almost all of the Treaty articles identified by the Provisional Government. In doing so Gavan Duffy tried the patience of a Provisional Government that was trying to get the best possible deal from the British Government while resisting armed opponents at home. Gavan Duffy's stance was close to the position that had been adopted by the Provisional Government and its advisors before the British–Irish negotiations on the draft Constitution. He had privately criticised the performance of Irish negotiators as a cabinet minister. Now, having resigned his cabinet post and sitting as an independent TD, he was free to give full rein to his conviction that the Irish negotiators had conceded more on the contents of the draft Constitution than was required by the terms of the 1921 Treaty. Gavan Duffy's position as a signatory of the Treaty and former minister ensured that he had to face a degree of hostility in his exchanges with his former colleagues, in particular W T Cosgrave, Kevin O'Higgins and Ernest Blythe, that none of the other critics of the Constitution had to endure. His attempts at amendment were criticised for their 'frightfully bad grammar'.[77] More seriously, he

76 A list of the 'Treaty articles' can be found in NAI, Department of the Taoiseach, S8956A, Kevin O'Higgins to Thomas Johnson, 22 September 1922. These provisions were sometimes called the 'vital clauses', for example *Dáil Debates*, vol 1, col 560, 21 September 1922 or the 'tied Articles', for example, *Dáil Debates*, vol 1, col 1084, 3 October 1922. Leo Kohn refers to them as the 'Agreed Articles', Kohn (n 62 above) 100.

77 *Dáil Debates*, vol 1, col 1669–1670, 18 October 1922.

was accused of running away from the responsibility he took on when signing the Treaty.[78] Kevin O'Higgins rejected one of his attempts to amend the draft Constitution by declaring that 'this Government will not dishonour the signature of Deputy Gavan Duffy, even at the invitation of Deputy Gavan Duffy'.[79] W T Cosgrave did not forebear from insisting that the Irish negotiators sent to London to negotiate key provisions of the Constitution had brought back a better draft than if Gavan Duffy had been a member of their team.[80] On one occasion it was even suggested that Gavan Duffy might try to form his own government to see if he could secure a better Constitution within the limits of the Treaty.[81]

The breakdown in relations between Gavan Duffy and his former colleagues was so complete that one opposition TD expressed surprise on one occasion in which government ministers appeared to agree with Gavan Duffy as 'they have always acted on the assumption that whatever Deputy Gavan Duffy favours must be wrong'.[82] Gavan Duffy did have to endure insults but proved fully capable of dealing out his own biting words. For example, he accused Kevin O'Higgins of 'schoolboy insolence'.[83] He condemned the 'poisoned fungus-growth which you will find all through this Constitution'.[84] Gavan Duffy accused the Provisional Government of 'abject surrender', deplored its 'weakness' and concluded 'It was no part of the Treaty that every time England should say "boo" to the Government the Government should run away.'[85] His conclusion that the draft Constitution was a 'surrender' and that the Irish delegation that had travelled to London had immediately given way to British demands were particularly bitter words from a person who had endured similar taunts with respect to his own involvement in the negotiations on the 1921 Treaty.[86]

It is also important to recognise that Gavan Duffy's involvement in the debates on the draft Constitution of the Irish Free State was not limited to the provisions that related to the 1921 Treaty that Gavan Duffy had signed some months earlier. He moved amendments on

78 *Dáil Debates*, vol 1, col 1006, 29 September 1922. See also Dáil Debates, vol 1, col 548, 21 September 1922.

79 *Dáil Debates*, vol 1, col 648, 25 September 1922.

80 *Dáil Debates*, vol 1, col 549, 21 September 1922.

81 *Dáil Debates*, vol 1, col 761, 26 September 1922.

82 *Dáil Debates*, vol 1, col 1541, 12 October 1922.

83 *Dáil Debates*, vol 1, col 997, 29 September 1922.

84 *Dáil Debates*, vol 1, col 538, 21 September 1922.

85 *Irish Independent* (n 60 above); and *Dáil Debates*, vol 1, col 532, 21 September 1922.

86 For example, one anti-Treaty newspaper referred to the document that Gavan Duffy had signed in 1921 as the 'Treaty of Surrender' *Republic of Ireland* (22 June 1922) 1.

topics as diverse as the definition of Irish citizenship,[87] the status of women,[88] courts martial,[89] Irish language titles[90] and on allowing members of one house of the Oireachtas to stand for election to the other.[91] Gavan Duffy also joined a special committee within the Constituent Assembly whose work resulted in a substantial redrafting of the provisions on the composition of the Executive.[92]

Yet, it was Gavan Duffy's interventions on the aspects of the draft Constitution that related to the 1921 Treaty that defined his political stance in 1922 and reflected the full extent of the breakdown in relations with former colleagues. Many of his attempts at amendment were based on an interpretation of the Treaty that permitted a synthesis of the law, practice and constitutional usage of Canada.[93] As mentioned earlier, the Provisional Government and its advisors had used a similar approach before the rejection of Draft D during the British–Irish negotiations on the draft Constitution.

The role and position of the King

Many Irish nationalists found it difficult to accept the reality that the future Irish Free State would come into existence as a constitutional monarchy. This included many supporters of the Treaty. Gavan Duffy objected to 'desecrating an Irish Constitution in the twentieth century with the royal relics of England's medievalism'.[94] Nevertheless, article 12 of the draft Constitution recognised the King as a constituent part of the Oireachtas alongside Dáil Éireann and Seanad Éireann. Kevin O'Higgins openly admitted that 'It is not a particularly pleasant position to be placed in to have to stand over an Article such as Article 12 of the Constitution.'[95]

Gavan Duffy attempted to use the 'synthesis' approach in an unsuccessful amendment that attempted to exclude the King from

87 *Dáil Debates*, vol 1, col 663–664, 25 September 1922.

88 For example, see *Dáil Debates*, vol 1, col 1681, 18 October 1922. Gavan Duffy felt that the draft gender equality clause needed to be redrafted and merited inclusion in a separate article. See NAI, Constitution Committee (n 48 above). See generally, Thomas Mohr, 'The rights of women under the Constitution of the Irish Free State' (2006) 41 Irish Jurist 20.

89 *Dáil Debates*, vol 1, col 1420, 10 October 1922.

90 *Dáil Debates*, vol 1, col 1395–1396, 10 October 1922.

91 *Dáil Debates*, vol 1, col 1038, 3 October 1922. This amendment influenced the final wording of article 16 of the 1922 Constitution which has, in turn, influenced article 15.14 of the 1937 Constitution of Ireland.

92 *Dáil Debates*, vol 1, col 1535, 12 October 1922.

93 *Dáil Debates*, vol 1, col 533–534, 21 September 1922 and col 761–81, 26 September 1922. See also *Irish Independent* (n 60 above).

94 Ibid.

95 *Dáil Debates*, vol 1, col 760, 26 September 1922.

the Oireachtas.[96] His efforts were unsuccessful as this provision was identified as one of the Treaty articles on which the Provisional Government would either stand or fall. O'Higgins described his proposed amendment as 'practically equivalent to an invitation to this Dáil to proclaim a Republic'.[97] Dominion precedent was clear and unambiguous on such matters and the Constitutions of Canada, Australia and South Africa were cited to defeat this amendment.[98] In a more light-hearted vein, a number of Labour representatives argued that since article 12 of the draft Constitution provided that the King was to be a member of the Oireachtas he should be required to travel to Dublin in order that he might make an 'oath of allegiance' to himself.[99]

Gavan Duffy also moved an unsuccessful amendment to remove the King from article 51 which recognised that 'The Executive Authority of the Irish Free State (Saorstát Éireann) is hereby declared to be vested in the King.'[100] He was also deeply concerned by a provision in article 83 under which the Constitution would come into operation on the issue of a 'Proclamation of His Majesty' not later than 6 December 1922. He recognised that it could be used to argue that the Constitution of the Irish Free State was a 'gift from England'.[101]

Gavan Duffy's attempts to remove the King from key parts of the 1922 Constitution met with a predictably hostile response from the Provisional Government. W T Cosgrave explained that there was a 'sprinkling' of references to the King in the Constitution because he also appeared in the Treaty that Gavan Duffy had signed. He added that Gavan Duffy may have been 'in his pyjamas and did not read the Article before he signed it'.[102]

The identity of the Irish Free State as a constitutional monarchy was always a difficult matter.[103] This would later be reflected in the total exclusion of monarchical symbols from the stamps and coins of

96 *Dáil Debates*, vol 1, col 762–763, 26 September 1922.

97 *Dáil Debates*, vol 1, col 759, 26 September 1922.

98 *Dáil Debates*, vol 1, col 759–760, 26 September 1922. See British North America Act 1867, s 17, article 1 of the 1900 Australian Constitution and South Africa Act 1909, s 19.

99 *Dáil Debates*, vol 1, col 1067–1072, 3 October 1922 and *Irish Independent* (n 60 above). Such comments indicate the confusion between the concepts of 'King' and 'Crown' which was common in the Constituent Assembly. This is not surprising when it is observed that the two terms are used interchangeably in the Irish Free State Constitution. See Barra O Briain, *The Irish Constitution* (Talbot Press 1929) 81 and 94.

100 See British North America Act 1867, s 9, article 61 of the 1900 Australian Constitution and South Africa Act 1909, s 8.

101 *Dáil Debates*, vol 1, col 1463, 11 October 1922.

102 *Dáil Debates*, vol 1, col 773–774, 26 September 1922.

103 For example, see Kohn (n 62 above) 114, 179 and 263.

the Irish Free State that were issued in the 1920s.[104] The references to the King in the Constitution were largely removed by the Constitution (Amendment No 27) Act 1936 and completed when the 1922 Constitution was replaced in 1937.

The representative of the Crown

Gavan Duffy was never comfortable with the representative of the Crown in Ireland being called the 'Governor-General' following the precedent of the existing Dominions such as a Canada. He objected to the name 'because it connotes the idea of domination'.[105] As a cabinet minister Gavan Duffy had proposed the 'British Commissioner' or the slightly sinister alternative of 'the British Agent'.[106] He argued in the Constituent Assembly that there was no necessity to put any name into the Constitution and that the matter might be left for determination at a future date.[107] Nevertheless, Ernest Blythe argued that this was an 'agreed clause'. He accused Gavan Duffy of making unfortunate jokes that the title should be 'An tAmadán Mór' (The Great Fool).[108] Accepting the term used in the existing Dominions was seen as increasing the chances of solidarity with the Irish Free State if the office were abused by future British governments.[109]

Gavan Duffy also made unsuccessful attempts to amend the provisions of the Constitution concerning the appointment of the Governor-General.[110] He also attempted to amend article 24 of the Constitution which concerned the summoning and dissolving of the Oireachtas by the Governor-General in the name of the King. Gavan Duffy wished to add words that clarified that this could only be done 'on the advice of the Executive Council'. This was deemed unnecessary by the Provisional Government and defeated in the Constituent Assembly.[111] Gavan Duffy also made unsuccessful efforts to amend

104 See Thomas Mohr, 'The political significance of the coinage of the Irish Free State' (2015) 23(4) Irish Studies Review 451–479.

105 Dáil Debates, vol 1, col 1770, 19 October 1922.

106 UCD Archives, Gavan Duffy Papers, P152/209, Gavan Duffy to Collins, 14 June 1922. He later proposed 'Commissioner of the British Commonwealth' in NLI, George Gavan Duffy Papers, MS 15,440/2/40, Orders of the day, 5 October 1922.

107 *Dáil Debates*, vol 1, col 1770–1771, 19 October 1922.

108 *Dáil Debates*, vol 1, col 773–774, 26 September 1922.

109 *Dáil Debates*, vol 1, col 1772–1773, 19 October 1922.

110 *Dáil Debates*, vol 1, col 1776–1779, 19 October 1922.

111 *Dáil Debates*, vol 1, col 1779–1783, 19 October 1922. See also NLI, George Gavan Duffy Papers, Ms 15,440/2/24, Orders of the Day, 25 September 1922.

constitutional provisions concerning the salary of the Governor-General which he considered excessive.[112]

The parliamentary oath

The controversial parliamentary oath that would be enshrined in article 17 of the Constitution was never destined to receive an enthusiastic reception in the Constituent Assembly. Gavan Duffy described its inclusion as 'abominable' and as 'one of the outstanding defects' of the draft Constitution.[113] Kevin O'Higgins admitted that 'it is not a pleasing task to stand over it, and it is not a pleasant task to submit it here to an Irish Assembly'.[114] The Provisional Government stuck to its well-established justification for the existence of the oath within the provisions of the Treaty. Article 2 of the Treaty implied that members of an Irish Parliament would have to swear an oath of allegiance to the King in the same manner as their Canadian counterparts. The Provisional Government stressed that the Irish had improved on the Canadian position in article 4 of the Treaty by securing a wording that pledged fidelity and allegiance to the Irish Constitution and fidelity to the King in a secondary capacity. [115]

Gavan Duffy noted that Article 4 of the Treaty simply provided that 'The oath to be taken by Members of the Parliament of the Irish Free State shall be in the following form …' and then detailed the form of the oath. He argued that these words, taken by themselves, did not place any legal obligation on members of the Oireachtas to take the oath. The line of reasoning used by Gavan Duffy was often called the 'whiskey argument'. It received this unusual soubriquet from the popular explanation of the proposition that the oath was optional for members of the Oireachtas. Exponents of this argument maintained that the provision in article 4 of the Treaty as to the form of oath to be taken by members of the Oireachtas was similar to a rule saying that if the members of a certain club wished to take whiskey it had to be 'John Jameson's Three Star'. However, the rules of the club did not forbid teetotalism and no rule could be found requiring members to drink whiskey. In the same way, it was argued that no clause of the Treaty required all members of the Oireachtas to take the oath. The Treaty

112 Article 60 of the Constitution of the Irish Free State. For example, see Dáil Debates, vol 1, col 1622–1623, 12 October 1922.

113 UCD Archives, Gavan Duffy Papers, P152/209, Gavan Duffy to Collins, 14 June 1922. *Irish Independent* (n 60 above). An earlier draft described the inclusion of the oath as 'the outstanding defect' in the draft Constitution. UCD Archives, Gavan Duffy Papers, P152/213, memorandum, 'Does the draft Constitution surrender more than the Treaty?', undated.

114 *Dáil Debates*, vol 1, col 1041, 3 October 1922.

115 *Dáil Debates*, vol 1, col 1039, 3 October 1922. See also Dáil Debates, vol 3, col 416–417, 25 April 1923.

merely provided that if they chose to take an oath it had to be in the form set out in article 4.[116]

George Gavan Duffy made full use of his position as a signatory of the Treaty to argue that the British Government had not intended the oath to be mandatory on all members of the Oireachtas. There were few remaining Irish signatories of the Treaty to contradict him on this point. Death had robbed the Provisional Government of Arthur Griffith and Michael Collins while a change in convictions ensured that Robert Barton was unavailable. Eamonn Duggan remained as the only signatory of the Treaty available to the Provisional Government. He was brought into the Constituent Assembly to refute his former colleague and stifle this dangerous amendment. Duggan assured the Constituent Assembly that all parties to the Treaty had been well aware that the oath was intended to be obligatory. It was this reality that had necessitated the lengthy and difficult negotiations as to the wording of the oath. Duggan added that he was in a better position to speak as to the nature of the oath since he, unlike Gavan Duffy, had been present at many of the conferences that had dealt with this matter.[117]

Despite the emphasis placed by Gavan Duffy on his position as Treaty signatory and as a former government minister, his attempts to circumvent the parliamentary oath were unsuccessful. He hoped that his initiative would 're-stabilise the country' even though the Civil War had already broken out by this date.[118] It is worth noting that anti-Treaty TDs had refused to join the debates of the Constituent Assembly which did not impose any parliamentary oath as a condition of attendance.[119] Nevertheless, in 1927 Gavan Duffy revived these arguments in advising de Valera that the parliamentary oath was not mandatory.[120] The 'whiskey argument' was raised again and again after 1922 including the period immediately after 1932 when de Valera came to power.[121] It was raised by Seán MacEntee, Minister for Finance, and by Conor Maguire, Attorney General, during the parliamentary debates on draft legislation that sought to remove the parliamentary oath from the text of the Constitution.[122] Abolition of the parliamentary oath finally took place with the enactment of the Constitution (Removal of Oath) Act 1933.

---

116 UCD Archives, Gavan Duffy Papers, P152/213, memorandum (n 113 above) and *Irish Independent* (n 60 above).

117 *Dáil Debates*, vol 1, col 1055–1056, 3 October 1922.

118 *Irish Independent* (n 60 above).

119 Ibid.

120 C Gavan Duffy (n 5 above) 96 and reprint 12.

121 For example, see Diarmuid Ó Crudhlaoich, *The Oath of Allegiance* (Maunsel & Roberts 1925) 67–93.

122 *Dáil Debates*, vol 41, col 591–596, 27 April 1932 and col 1010, 29 April 1932.

The power of reservation

Article 41 of the draft Constitution referred to an institution recognised throughout the Commonwealth as the power of 'reservation'. The power of reservation was initiated by the representative of the Crown in a Dominion who could withhold the King's assent to a Bill passed by a Dominion legislature pending the signification of the King's pleasure. It was a delaying measure that could be converted into a permanent veto.[123]

The Provisional Government had ensured that Draft D omitted all reference to the power of reservation. The British demanded its insertion as a Dominion institution during the Constitution negotiations of 1922.[124] The power to reserve legislation passed by the Oireachtas was finally recognised in article 41 of the Constitution.

As might be expected, the Provisional Government played down the importance of this feature of the Constitution. Kevin O'Higgins claimed that 'this is a nominal and theoretical veto' and assured the Constituent Assembly that there was no need to take article 41 at face value. He also told the Constituent Assembly that constitutional lawyers now recognised that the power of reservation could no longer be used with respect to legislation concerning internal matters.[125]

As might be expected, George Gavan Duffy proposed an amendment to article 41 during the debates of the Constituent Assembly. This amendment would have declared that the powers of veto in article 41 were obsolete in Canada and, by extension, were also obsolete with respect to the Irish Free State. He concluded: 'If the veto is dead let us say so.'[126] The Provisional Government could not accept such a fundamental alteration of one of the 'Treaty articles' of the Constitution.[127] O'Higgins was forced to admit that the power of reservation could still be exercised with respect to legislation that affected the United Kingdom or the existing Dominions.[128] In taking this position, O'Higgins was repeating the views pressed by Lloyd

123 On some occasions, the royal assent was withheld from a reserved Bill with the full agreement of the Government of the Dominion concerned. This was the case with respect to the Australian Customs Tariff (British Preference) Bill 1906 and the New Zealand Shipping and Seamen (Amendment) Bill 1910. K C Wheare, *The Statute of Westminster and Dominion Status* 4th edn (Oxford University Press 1949) 68–69.

124 During Kennedy and Hewart's redrafting of the Constitution on 6 to 9 June, a new provision, article 39A, was placed in the draft Constitution, which became article 40 of the Anglo Irish Draft and article 41 of the 1922 Constitution. TNA, CAB 43/2 SFB 63, Anglo Irish Draft Constitution.

125 *Dáil Debates*, vol 1, col 1168–1169, 4 October 1922.

126 *Dáil Debates*, vol 1, col 1172, 4 October 1922.

127 *Dáil Debates*, vol 1, col 1171–1172 and 1183, 4 October 1922.

128 *Dáil Debates*, vol 1, col 1168–1169, 4 October 1922.

George and the British law officers during the negotiations on the 1922 Constitution.[129] Once again, Gavan Duffy was scratching at a sore point that caused serious embarrassment to the Provisional Government. The power of reservation was formally abolished by the Constitution (Amendment No 27) Act 1936.

Popular sovereignty

Article 2 of the draft Constitution, as presented to the Constituent Assembly, provided that legislative, executive and judicial powers were derived from the people. This expression of popular sovereignty was an anomaly among the Treaty articles of the Irish Free State Constitution. The other Treaty articles were placed in the text at the insistence of a British Government that wanted to ensure that the Irish Free State had all the legal accoutrements of a British Dominion. By contrast, article 2 was presented as a Treaty article at the insistence of the Irish Provisional Government. The Provisional Government presented British acceptance of this provision as a significant negotiating victory.[130] The Provisional Government was determined to protect it from a Constituent Assembly that might not fully appreciate its value.

Article 2, as one of the Treaty articles, should not have been capable of amendment without threatening the continuance in office of the Provisional Government. Nevertheless, it was altered with the reluctant consent of Irish ministers. The provision presented to the Constituent Assembly provided that 'All powers of government and all authority legislative, executive and judicial, are derived from the people.' Gavan Duffy, in association with other TDs, argued that it was necessary to amend this provision to declare that these powers be derived from 'the *Irish* people'.[131] At first, the Provisional Government rejected this as a spurious amendment.[132] What other people could have been intended in the context of an Irish Constitution? Nevertheless, Gavan Duffy pointed out that the day would come when the provisions of the Irish Constitution would be interpreted by the Judicial Committee of the Privy Council. Fears were expressed as to the possibility that a mischievous Privy Council might interpret the term 'the people' as referring to the people of the Commonwealth, on the basis of their 'common citizenship', in place of the people of the Irish Free State

129 TNA, CAB 43/3 SFC 35, British memorandum on draft Irish Free State Constitution, 20 May 1922.

130 For example, see *Dáil Debates*, vol 1, col 655, 25 September 1922.

131 See UCD Archives, Kennedy Papers, P4/341, private and confidential, 18 September 1922.

132 See UCD Archives, Kennedy Papers, P4/340, Constitution, committee stage, notes, undated.

alone.[133] The demand for a specific reference to 'the Irish people' was later amended to 'the people of Ireland' on the basis that the former might have been interpreted to include all persons of Irish descent from around the globe.[134] The declaration of popular sovereignty that appeared in article 2 of the 1922 Constitution was reproduced with some amendments in article 6.1 of the 1937 Constitution of Ireland.

The legal supremacy of the treaty

As seen earlier, Gavan Duffy was opposed to any recognition that the Treaty he had signed in 1921 enjoyed force of law.[135] He wanted to unilaterally change the official name of the document from 'the Articles of Agreement for a Treaty between Great Britain and Ireland' to 'the Treaty between Great Britain and Ireland'.[136] This was part of his policy of keeping maximum distance between the Treaty that he signed and the Irish Constitution which he argued 'must be kept as distinct as possible'.[137]

In these circumstances there was little surprise that Gavan Duffy strongly objected to the provisions of the Constitution that would later be known as the Repugnancy Clause. The presence of the Repugnancy Clause, which asserted the supremacy of the 1921 Treaty over all sources of Irish law including the Constitution and amendments to it, had been a contributing factor to George Gavan Duffy's break with the Provisional Government.[138] He argued that 'It makes the Treaty a part of the Constitution and it seems to make the Constitution derive its force from the Treaty.'[139] Gavan Duffy also believed that 'it has the air of making our right of freedom of government depend upon and derive from an Agreement with England and gives far too important a place to the Treaty'.[140] He was also keen to emphasise that 'There is nothing in

133 *Dáil Debates*, vol 1, col 655–656, 25 September 1922.

134 Ibid 661.

135 NAI, Constitution Committee (n 48 above).

136 NLI, George Gavan Duffy Papers, MS 15,440/2/94, Draft Constitution of Saorstát Éireann as amended in Committee. Gavan Duffy also wished to remove all reference to the 1921 Treaty in the long title of the Irish statute containing the 1922 Constitution. This was 'An Act to enact a Constitution for the Irish Free State (Saorstát Éireann) and for implementing the Treaty between Great Britain and Ireland signed at London on the 6th Day of December, 1921'. Gavan Duffy wanted to remove all words after (Saorstát Éireann). NLI, George Gavan Duffy Papers, MS 15,440/2/94, draft Constitution of Saorstát Éireann as amended in Committee.

137 *Irish Independent* (n 60 above).

138 NAI, Department of the Taoiseach, S8955, Gavan Duffy to Collins, 14 June 1924.

139 UCD Archives, Gavan Duffy Papers, P152/209, Gavan Duffy to Collins, 14 June 1922.

140 UCD Archives, Gavan Duffy Papers, P152/213, memorandum (n 113 above).

the Treaty to make it a compact for all time.'[141] Gavan Duffy appears to have anticipated the possibility of a new revised treaty at some point in the future that would replace the document he had signed. The Repugnancy Clause stood in the way of this evolution.[142] Gavan Duffy's final argument, which he made again and again, was that no provision of the Treaty that he had signed required acceptance of such a provision.[143]

Gavan Duffy made it abundantly clear that he wished to remove the Repugnancy Clause.[144] He believed it to be the worst aspect of the draft Constitution with the single exception of the provisions on the parliamentary oath.[145] Yet, when the Repugnancy Clause finally came before the Constituent Assembly, he made no effort to remove or amend it. The debates on this provision came near the end of the debates of the Constituent Assembly, and Gavan Duffy may have become bruised by the failure of most of his previous efforts. Instead of attempting amendment he simply declared that the Repugnancy Clause was 'a denial of our sovereignty, and is about as bad as it could be, and therefore I do not think we should touch a line of it, as it is a fitting introduction to the emaciated Constitution'.[146]

At this stage Gavan Duffy may have been resigned to defeat or believed that it might be wiser to support the attempts at amendment made by others which might have had a greater chance of success. He had considered moving an amendment to article 50 that would have created an unusual mechanism to amend the Constitution without any need for a referendum. This could be used to amend provisions of the Constitution that curtailed 'the natural rights of this free people further or otherwise than those rights shall have been curtailed by Treaty amendments'.[147] It should be recalled that Gavan Duffy did not believe that most of the Treaty articles reflected the strict terms of the Treaty. In the end, Gavan Duffy chose to withdraw this proposal which never had any real chance of acceptance.[148] In September 1922, some months after Gavan Duffy's resignation as cabinet minister,

141 *Irish Independent* (n 60 above).

142 UCD Archives, Gavan Duffy Papers, P152/213, memorandum (n 113 above) and *Irish Independent* (n 60 above).

143 *Dáil Debates*, vol 1, col 533, 21 September 1922 and *Irish Independent* (n 60 above).

144 For example, see NLI, George Gavan Duffy Papers, MS 15,440/2/93, Draft Constitution of Saorstát Éireann Bill 1922.

145 UCD Archives, Gavan Duffy Papers, P152/209, Gavan Duffy to Collins, 14 June 1922.

146 *Dáil Debates*, vol 1, col 1481, 11 October 1922.

147 NLI, George Gavan Duffy Papers, MS 15,440/2/40, Orders of the day, 5 October 1922.

148 *Dáil Debates*, vol 1, col 1239, 5 October 1922.

the Provisional Government had made its own efforts to remove the Repugnancy Clause during secret negotiations with the British Government but without success.[149] The text of the Repugnancy Clause, together with the corresponding provision in article 50, entered the Irish Constitution intact.[150]

Gavan Duffy feared that the Repugnancy Clause would impede advances in Irish autonomy. This may have underpinned his strong support for amendments to article 50 that made it possible to amend the Constitution without the need for a referendum for a period of eight years. The change would facilitate the removal of the Treaty articles that Gavan Duffy had attempted to resist. Yet, this proved to be a controversial change to the Constitution because, as events transpired, the period of eight years was substantially extended and, consequently, the referendum provisions never worked as intended. Nevertheless, there is little doubt that a flexible Constitution that could be easily amended by the Oireachtas would facilitate the swift removal of undesirable provisions at an opportune moment. It certainly facilitated the removal of the Repugnancy Clause, in tandem with the removal of the parliamentary oath, by de Valera's Fianna Fáil Government in 1933.[151]

The legislative supremacy of the Imperial Parliament

In 1922 the Parliament at Westminster was still popularly known as the 'Imperial Parliament'. The term did have some reality as this Parliament retained the power to legislate for any constituent part of the British Empire. Yet, by the early twentieth century it was recognised that, in practice, Westminster could only legislate for the self-governing Dominions with their consent. This reality was reflected in the deliberations of the Imperial War Conference of 1918 when several Dominion premiers emphasised this point.[152] Nevertheless, all the existing Dominion constitutions had come into existence as British 'Imperial' statutes passed at Westminster and the British

149 Thomas Towey, 'Hugh Kennedy and the constitutional development of the Irish Free State, 1922–1923' (1977) 13 Irish Jurist 355, 362-4.

150 It is worth noting that the wording of the Repugnancy Clause seemed to overlook the fact that, after a transitional period during which constitutional amendments could be made by the Oireachtas, a referendum was required under article 50 to alter the text of the Constitution. Although this point was initially of some concern to the British Government, no stipulation was placed in the Repugnancy Clause dealing with this potential difficulty. It is likely that they concluded that the effect of the Repugnancy Clause would override the need for holding a referendum. TNA, CAB 43/3 SFC 35, British memorandum on draft Irish Free State Constitution, 29 May 1922.

151 Constitution (Removal of Oath) Act 1933.

152 Henry Harrison, *Ireland and the British Empire, 1937* (Hale 1937) 148.

Government was determined that the same must apply to the Irish Free State. Consequently, the 1922 Irish Free State Constitution was enacted in two parallel statutes, one passed by the Dáil, sitting as a Constituent Assembly, in Dublin and the other by the Parliament at Westminster. This reality of dual creation was recognised by article 83 of the 1922 Constitution itself. Gavan Duffy moved an amendment in the Constituent Assembly that sought to downgrade the British statute and emphasise the supremacy of the one passed in Dublin but, once again, the Provisional Government used its parliamentary support to defeat a challenge to this Treaty article.[153]

Gavan Duffy was also concerned at the prospect that the Imperial Parliament might attempt to pass additional legislation for the Irish Free State after it officially came into existence on 6 December 1922. This remained possible under the law of the time, and Imperial statutes actually enjoyed superior status over those passed by a Dominion parliament, a reality recognised by the Colonial Laws Validity Act 1865.[154] Once again, Gavan Duffy moved an amendment to the draft Constitution based on an argument reliant on a synthesis of 'law, practice and constitutional usage'.[155] However, on this occasion Gavan Duffy's argument was fortified by a letter written to *The Times* by Arthur Berriedale Keith, professor at the University of Edinburgh and a leading authority on British Imperial law.[156] He emphasised that Keith was 'a true blue Briton and not an Irish Rebel' and yet had recommended the same change now being championed by Gavan Duffy that would emphasise Irish legislative autonomy.[157] Gavan Duffy moved to amend the concluding sentence of article 12 of the draft Constitution by adding the words 'sole and exclusive' in order that it would read:

> The *sole and exclusive* power of making laws for the peace, order and good government of the Irish Free State (Saorstát Éireann) is vested in the Oireachtas.[158]

153 Article 83 of the 1922 Constitution of the Irish Free State provided for 'The passing and adoption of this Constitution by the constituent assembly and the British Parliament.' Gavan Duffy wished to replace the words 'passing and adoption' with 'registered' in relation to the British Parliament. *Dáil Debates*, vol 1, col 1458–1462, 11 October 1922.

154 Colonial Laws Validity Act 1865, s 2.

155 *Dáil Debates*, vol 1, col 778–780, 26 September 1922.

156 *Dáil Debates*, vol 1, col 779–780, 26 September 1922.

157 *Dáil Debates*, vol 1, col 779, 26 September 1922. Gavan Duffy also emphasised Keith's letter in a newspaper article in which he wrote 'Fas est et ab hoste doceri' or 'it is right to be taught even by an enemy'. *Irish Independent* (n 60 above).

158 Italicised words added by George Gavan Duffy. *Dáil Debates*, vol 1, col 777–781, 26 September 1922.

On this occasion the Provisional Government was actually prepared to consider George Gavan Duffy's proposal. It was true that article 12 was one of the specified Treaty articles, and the Provisional Government did not seem inclined, at first, to accept Gavan Duffy's proposal.[159] In the end, the Provisional Government relented and accepted Gavan Duffy's amendment to article 12.

This amendment, the only significant change made by the Constituent Assembly to the Treaty articles of the Constitution, did not go unnoticed in London. The British Government considered the amended version of article 12 to be a breach of the 1921 Treaty but was unwilling to risk the entire settlement by insisting on its amendment at the eleventh hour. Instead, it decided to add additional provisions to the text of the British statute establishing the Irish Constitution. These included a provision intended to safeguard the right to pass Imperial legislation for the Irish Free State on matters of common concern, as was the case in the other Dominions, while leaving intact the Irish monopoly on legislation affecting the 'peace, order and good government' of the Irish Free State.[160] The power saved by this provision was actually used on a number of occasions after 1922.[161]

Gavan Duffy's amendment, in referring to the 'sole and exclusive power' of the Oireachtas to make laws for the Irish state would be reproduced in article 15.2.1 of the 1937 Constitution of Ireland. However, with the passage of time the original context of the insertion of these words was largely forgotten. Instead, they provided the basis of extensive case law on the doctrine of separation of powers.[162] Gavan Duffy's amendment proved to be of considerable significance in ways that could not be fully appreciated in 1922.

The Privy Council appeal

Ironically, one of the few Treaty articles that Gavan Duffy did not contest in the Constituent Assembly proved, in practice, to be most threatening to Irish autonomy. The provisions for the appeal to the Judicial Committee of the Privy Council, the supreme appellate court

159 Kevin O'Higgins argued that the presence of the definite article before the word 'power' was sufficient to show that the Oireachtas already had sole and exclusive power to legislate for the Irish Free State. He also pointed out that article 2 of the Constitution provided that all powers within the Irish Free State had to be exercised by organisations established by or under and in accord with that Constitution which clearly did not apply to the Parliament at Westminster. *Dáil Debate*s, vol 1, col 780–781, 26 September 1922.

160 Irish Free State Constitution Act 1922 (Session 2), s 4.

161 See Thomas Mohr, 'British Imperial statutes and Irish sovereignty: statutes passed after the creation of the Irish Free State' (2011) 32(1) Journal of Legal History 61.

162 For example, see *Cityview Press Ltd v An Chomhairle Oiliúna* [1980] IR 381.

for most of the British Empire, would appear in article 66 of the Constitution. The Privy Council appeal was debated towards the end of the proceedings of the Constituent Assembly and, as seen earlier, Gavan Duffy may have realised that amendments moved by him had little chance of acceptance and preferred to support attempts made by other parties. In any case, Gavan Duffy was less concerned with this provision than with other Treaty articles of the 1922 Constitution even though he did conclude that it was an insult to Irish dignity.[163] He explained that the reason he had decided not to contest the Privy Council appeal was his conviction that it would be a 'dead letter'.[164] He concluded that 'this appeal is a humbug from beginning to end' and predicted that there would never be an appeal from the Irish courts to the Judicial Committee of the Privy Council.[165] Gavan Duffy's confidence on this matter proved misplaced. The Irish appeal to the Privy Council did function, and proved to be one of the most controversial aspects of the 1922 Constitution until its final abolition in the mid-1930s. Despite Gavan Duffy's legal experience, he never fully appreciated the British Government's intention that the Privy Council serve as arbiter of the Treaty settlement.[166]

ALTERNATIVES TO COMPROMISE ON THE 1922 CONSTITUTION

Gavan Duffy's strong stance on the Treaty articles of the 1922 Constitution could be blamed on his conviction that British Imperial law, as it related to the Dominions, would soon be radically overhauled. He was convinced that an Imperial Constitutional Conference, anticipated in the published deliberations of the Imperial War Conference of 1917, was just around the corner. This was an important consideration in convincing Gavan Duffy that the legal provisions derived from Dominions were being enshrined in the 1922 Irish Constitution at the worst possible moment.[167]

Gavan Duffy focused on a decision, known as Resolution IX, made by the British and Dominion Prime Ministers attending the Imperial War Conference of 1917. This resolution recognised that a time of war was not a suitable occasion for the readjustment of the constitutional

163 *Dáil Debates*, vol 1, col 1413, 10 October 1922. Gavan Duffy appeared to class the Privy Council appeal with more minor consequences of the 1921 Treaty. See UCD Archives, Gavan Duffy Papers, P152/213, memorandum (n 113 above).

164 *Dáil Debates*, vol 1, col 1413, 10 October 1922.

165 *Dáil Debates*, vol 1, col 1414, 10 October 1922.

166 See Thomas Mohr, *Guardian of the Treaty – The Privy Council Appeal and Irish Sovereignty* (Four Courts Press 2016).

167 *Dáil Debates*, vol 1, col 538 and 542–543, 21 September 1922.

relations of the component parts of the Empire. However, the resolution added that such a readjustment 'should form the subject of a special Imperial Conference to be summoned as soon as possible after the cessation of hostilities'.[168] Resolution IX also provided that any such readjustment 'should be based upon a full recognition of the Dominions as autonomous nations of an Imperial Commonwealth'.[169]

Gavan Duffy advocated an immediate convocation of the anticipated Imperial Constitutional Conference when the British Government rejected Draft D of the future Constitution of the Irish Free State in May 1922. He was convinced that the convening of this anticipated conference would be 'immensely popular in the Dominions' and that it would resolve all of the major points in dispute in the draft Irish Constitution.[170] This argument lost its force when Gavan Duffy was corrected in the press and in the Constituent Assembly on his claims that an Imperial Constitutional Conference remained imminent.[171] Attempts at constitutional reforms in the direction of greater Dominion autonomy had actually been defeated at the Imperial Conference of 1921. Moreover the report of that conference appeared to close the door on future attempts at securing constitutional reform when it concluded that 'no advantage is to be gained by holding a constitutional Conference'.[172] The immediate prospect of an Imperial Constitutional Conference that would usher in sweeping changes had collapsed even before Gavan Duffy had signed the Treaty.

The collapse of prospects for a special Imperial Constitutional Conference had little impact on Gavan Duffy's stance on the Constitution. He was convinced that the Provisional Government had failed to consider other viable options to immediate compromise with the British Government on the draft Constitution. These included invoking the compulsory arbitration of the League of Nations.[173] This ignored the reality that the embryonic Irish Free State was not a member of the League in mid-1922 and, in the eyes of the international community, remained a part of the United Kingdom until 6 December 1922. His alternative recommendation of an appeal to the Dominions

168 Parliamentary Papers, 1917, Cmd 8566, Imperial War Conference 1917, 5 and 60.

169 Ibid.

170 UCD Archives, Gavan Duffy Papers, P152/206, minutes of a Cabinet meeting, 5 June 1922 and P152/207, Gavan Duffy to Collins, 3 June 1922.

171 *Freeman's Journal* (Dublin, 23 September 1922) 6 and *Dáil Debates*, vol 1, col 1176, 4 October 1922.

172 Cmd 1474, 9, Resolution XIV. See also TNA, CAB 32/6, Imperial Meetings 1921, p 323.

173 *Irish Independent* (n 60 above); and George Gavan Duffy, 'Notice sur la Constitution de l'État libre d'Irlande' (1922) 20 Annuaire de Legislation Étrangère 180.

was also hampered by the reality that the Irish Free State had not yet joined their ranks at this time. He may, in any case, have underestimated the difficulty in convincing Dominion statesmen that the Irish Free State should not have to accept institutions that were recognised in all of their Constitutions.

Gavan Duffy's final option of having the Provisional Government resign and leaving the British Government 'face to face with Mr de Valera and the opponents of the Treaty' raised a real prospect of renewed war.[174] In 1951 Gavan Duffy would reiterate his conviction that the British Government would have given way on the draft Constitution in the summer of 1922 as 'to denounce the much vaunted Anglo–Irish Treaty of Peace would have been too humiliating a solution'.[175] Although Gavan Duffy's conclusion is open to challenge, it is clear that the events of 1922 continued to preoccupy him until the end of his life.

ENGAGEMENT WITH THE OVERSEAS DOMINIONS

Despite Gavan Duffy's disappointment on the prospect of immediate constitutional reform, he remained a strong advocate of cooperation with the overseas Dominions at future Imperial Conferences. He remained convinced that these conferences offered much potential in advancing the autonomy of the Irish Free State. This proved to be one of Gavan Duffy's most insightful contributions after the signing of the Treaty. This route to constitutional reform was not nearly as obvious to contemporaries as it would be to subsequent commentators enjoying the benefit of hindsight.[176] It was openly rejected by many opponents and supporters of the 1921 Treaty who could not be sure that the restless Irish Free State would find any friends among the overseas Dominions. Gavan Duffy's argument was not helped by serious mistakes and misconceptions, in particular his belief that an Imperial Conference dedicated to constitutional reform was due to be convened in the near future. Yet, although Gavan Duffy was clearly mistaken in terms of details, his instinct as to the direction of reform and the prospect of finding common cause with other Dominions proved to be entirely accurate.

There was little precedent for the Imperial Conferences as a venue for seeking constitutional reform in the early 1920s. This would change as the decade wore on. The Imperial Conferences of later 1920s and

174 Gavan Duffy (n 173 above) 183. Translation provided in Dorothy Macardle, *The Irish Republic* (Irish Press 1951) 725.

175 See Military Archives (n 7 above).

176 Gavan Duffy's son Colum Gavan Duffy would later argue that his father's legal arguments before the constituent assembly 'were made before their time'. C Gavan Duffy (n 5 above) 95 and reprint 10.

early 1930s would culminate in the historic Statute of Westminster Act 1931. The reforms ushered in by this celebrated statute provided the bedrock for the constitutional reforms initiated by de Valera in which Gavan Duffy would also play a role. These would gradually dismantle the settlement initiated by the 1921 Treaty that Gavan Duffy had never fully accepted. The new Constitution of Ireland adopted in 1937 had no legal basis in the Treaty settlement. A British–Irish agreement in 1938 would dismantle the defence provisions of that settlement.[177] The Irish state would finally become a republic and leave the Commonwealth in 1949. Gavan Duffy's instincts on the potential offered by this peaceful avenue for constitutional change would, in the long term, prove to be justified.

CONCLUSION

George Gavan Duffy lost his parliamentary seat in the 1923 general election. This may have been a consequence of perceptions that Gavan Duffy had fallen between two stools. His position as a signatory of the Treaty made him unacceptable to opponents of the settlement while his stance in the Constituent Assembly and the bitter clashes with the Provisional Government may have rendered him unattractive to supporters of the Treaty. Gavan Duffy would later condemn the state of public opinion at the time as taking 'a childlike view of politics' that perceived support or opposition to the Treaty in terms of 'pure white' and 'murky black'. He concluded that anyone who was not 'wholeheartedly with the white or the black' was considered a 'crank', a 'factionist' or a 'Tadhg an dá thaobh' (someone who tries to satisfy both sides).[178] Gavan Duffy's son, Colum Gavan Duffy, would later conclude that it may have been as well that his father lost his parliamentary seat in 1923 as 'his scholarly temperament and his notion of absolute integrity would not have suited him to be an active and successful politician'.[179] Gavan Duffy would attempt to win a seat in the Seanad in the early 1930s but this proved unsuccessful.[180]

Gavan Duffy had never been comfortable with his position as a cabinet minister, and his resignation left him an isolated figure whose future in politics was precarious. He remained, as a former cabinet minister, a source of embarrassment to the Irish Government in the Constituent Assembly and afterwards. He did not shy away from revealing that the draft Constitution that had travelled to London had

177 See Eire (Confirmation of Agreements) Act 1938 (UK).

178 See Military Archives (n 7 above). 'Tadhg an dá thaobh' can be literally translated as 'Timothy of both sides'.

179 C Gavan Duffy (n 5 above) 95 and reprint 10.

180 Ibid 96–97 and reprint 12.

been a very different document to the draft that had returned even though this was a strictly confidential matter. His electoral leaflets for the 1923 election declared that the Government 'threw away in London a splendid Constitution drawn up by their own experts – a Constitution which would have gone a very long way to secure peace in Ireland – and substituted for it a Constitution mangled to the orders of Downing Street'.[181] Gavan Duffy would later declare that Griffith 'at once gave way' when faced with British pressure and condemned the resulting negotiations as a 'surrender'.[182]

Gavan Duffy's poor opinion of the performance of the Irish delegation during the British–Irish negotiations on the draft Constitution underestimate the serious pressure that they faced. His allegations of 'surrender' also provided little room for appreciating the negotiating successes achieved at this conference. These included the recognition of Irish popular sovereignty that appeared in article 2 of the Constitution. The inclusion of this provision was particularly significant as opponents of the Treaty had predicted that the British Government would never concede such a provision.[183] The Irish delegation that had travelled to London in late 1921, including Gavan Duffy, had tried and failed to include such a provision in the text of the Treaty.[184]

Gavan Duffy had a poor opinion of the Constituent Assembly of 1922 and would later recount that 'Only a handful of men made any serious contribution to the debates.' He regretted the absence of anti-Treaty TDs whose presence, he was convinced, would have resulted in 'a much improved Constitution'. In their absence, Gavan Duffy could do no more than accuse pro-Treaty Sinn Féin (which would soon become Cumann na nGaedheal) for its 'amazing docility' which had left the Provisional Government 'secure in the solid caucus behind it'.[185] Yet, this party also displayed remarkable solidarity in the context of making unpopular decisions that must also have been personally distasteful to many TDs. Gavan Duffy had been frustrated by an early example of party discipline in the politics of the developing Irish state.[186]

---

181 NLI, George Gavan Duffy, *Is it Fair?* (Dublin, 1923) Call Number: EPH C116.

182 See Military Archives (n 7 above).

183 *Clause by Clause – A Comparison between the 'Treaty' and Document No 2, Ireland* (Republic of Ireland 1922) 5.

184 Ibid.

185 See Military Archives (n 7 above).

186 For a contrasting view, see Laura Cahillane, *Drafting the Irish Free State Constitution* (Manchester University Press 2016) 75. A small number of TDs who identified with pro-Treaty Sinn Féin, for example Eoin MacNeill and Liam de Róiste, took an independent line, but overall party discipline in the face of the unpopular Treaty articles was overwhelming.

Gavan Duffy challenged the Provisional Government to publish Draft D that had been taken to London by Irish negotiators.[187] He even made thinly veiled threats that he would publish Draft D himself but warnings from the Provisional Government against breaches of confidence appear to have dissuaded him.[188] Ernest Blythe insisted that publication would set a bad precedent and added that Gavan Duffy had a purely personal motive in urging publication in allowing him 'to run away from the responsibility which he took when he signed the Treaty, and which he has since consistently tried to run away from'.[189] Gavan Duffy was accused of naivety in failing to appreciate that the Irish negotiators who went to London would have to make some concessions on the draft Constitution.[190]

The conclusion of his political career saw Gavan Duffy return to work at the bar but political developments soon opened new opportunities. In 1926 Eamon de Valera and his followers left Sinn Féin and formed a new political party known as 'Fianna Fáil'. The following year saw de Valera and his colleagues take the controversial parliamentary oath and enter the Oireachtas. De Valera now had a common goal with Gavan Duffy in seeking to push the legal interpretation of the Treaty settlement to its limits as a prelude to its dismantling. In 1927 Gavan Duffy publicly proclaimed that there was no legal obligation to continue paying land annuities to the United Kingdom and also advised de Valera that the parliamentary oath provided in the 1921 Treaty was not mandatory.[191] When de Valera came to power in 1932 he proposed to appoint Gavan Duffy as attorney general but, in the end, declined to do so due to internal opposition within his own Cabinet based on Gavan Duffy's status as a signatory of the Treaty.[192] Nevertheless, the close relationship that developed with de Valera is likely to have influenced Gavan Duffy's appointment to the High Court in 1936.

Gavan Duffy's political career was relatively short-lived and fell between two periods in the legal professions that proved far more successful. Nevertheless, Gavan Duffy's career between the signing of the Treaty in 1921 and the loss of his parliamentary seat in 1923 reveals much about him. One obvious feature was his isolation as a

187 *Dáil Debates*, vol 1, col 762, 26 September 1922.

188 *Dáil Debates*, vol 1, col 641–642, 25 September 1922.

189 *Dáil Debates*, vol 1, col 1006, 29 September 1922. Gavan Duffy would publish an article outlining some details of Draft D but, as it appeared in the French language in a legal periodical published in Paris, it is questionable how many contemporary Irish readers ever accessed it. Gavan Duffy (n 173 above).

190 *Dáil Debates*, vol 1, col 1005, 29 September 1922.

191 C Gavan Duffy (n 5 above) 96 and reprint 12.

192 Ibid 97–98 and reprint 14. Gavan Duffy would also come to admire de Valera's 'Document No 2' an alternative to the 1921 Treaty that Gavan Duffy had signed. See Military Archives (n 7 above).

government minister that was evident long before his final resignation. This isolation became even more pronounced in his brief career as an independent TD. He was accused of delusion and of trying to convince the Constituent Assembly that the Treaty had provided for a republic.[193] Gavan Duffy was also accused of treating the 1921 Treaty like a 'curate's egg' which, being 'excellent in parts', allowed him to pick and choose parts of the Treaty that he liked to the exclusion of those he did not.[194] On other occasions he was simply accused of attempting to 'run away' from the Treaty that he himself had signed.[195]

It should be recognised that Gavan Duffy acknowledged that the 1922 Constitution was an important advance for Irish nationalism notwithstanding his conviction it had failed to grasp the full degree of status and autonomy available under the 1921 Treaty.[196] The sincerity of Gavan Duffy's position in 1922 is reflected in the fact that he maintained it until the end of his life. The proceedings of the Constituent Assembly touched him deeply, and he concluded a written account of the Constituent Assembly's acceptance of the Treaty articles with the words 'Hinc illae lacrymae' (hence those tears).[197] Nevertheless, the debates of the Constituent Assembly suggest that he had a flawed understanding of the legal and political consequences of the Treaty that he signed in 1921 with such reluctance. It is also difficult to escape the conclusion that he maintained a deep sense of uneasiness, perhaps even regret, over this settlement. This is not surprising as Gavan Duffy proved to be the last of the Irish delegation to agree to sign the Treaty. It is also worth noting his subsequent justification for the final decision to sign which focused on arguments relating to duress during the last stage of the negotiations coupled with hopes that the drafting of the Constitution of the Irish Free State might mitigate some of the most objectionable features of the Treaty settlement. Gavan Duffy was not alone in hoping that the Constitution would offer a 'second round' in the negotiation of the British–Irish settlement, but he was alone among Irish cabinet ministers in refusing to accept the failure of this stance in the British–Irish negotiations of mid-1922 that produced an agreed draft. Gavan Duffy appears to have treated the deliberations of the Constituent Assembly in late 1922 as a 'third round' in negotiating a settlement. His anticipated Imperial Constitutional Conference, which was never based in reality, even appeared to have offered a 'fourth round'. His opposition to the Repugnancy Clause revealed his

193 *Dáil Debates*, vol 1, col 573, 21 September 1922.
194 Ibid.
195 *Dáil Debates*, vol 1, col 1006, 29 September 1922.
196 *Dáil Debates*, vol 1, col 1913–1914, 25 October 1922.
197 Gavan Duffy (n 173 above) 183.

conviction that the 1921 Treaty might be replaced in the near future by a revised British–Irish treaty which might be seen as a 'fifth round'.[198]

Gavan Duffy's legal arguments in the Constituent Assembly were based on a sincere desire 'to save something from the wreck of the Constitution on the rocks of Downing Street'.[199] Nevertheless, they were certainly open to challenge and he championed positions that had already been rejected twice, during the negotiations on the Treaty in 1921 and during the negotiations on the Constitution in 1922. He also persisted in a belief, evident during the negotiations on the Treaty and on the Constitution, that the British Government would be prepared to concede a self-governing Irish state with far more autonomy than Dominion status in the early 1920s.

Gavan Duffy was convinced that the British position on the draft Constitution was based on the 'bad example' it would set for nationalists within the Dominions.[200] He may not have fully appreciated that British resistance had much deeper roots. The British Government needed the world, and its own restless colonies, to see that, although the territory of the future Irish Free State was going to secede from the United Kingdom, it was not going to secede from the British Empire. If the Irish Free State was to look like a Dominion, its Constitution would have to make reference to key institutions that were also mentioned in the Constitutions of all the existing Dominions.

Gavan Duffy's attitude to the draft Constitution was broadly similar to that of his cabinet colleagues before the draft Constitution was taken to London in May 1922. It was his excessively optimistic perspective on what the British Government might be prepared to accept that set him apart during the British–Irish negotiations that followed. He was also convinced that the Provisional Government had not considered alternatives to compromising with the British Government. These included the convocation of an anticipated Imperial Constitutional Conference and arbitration by the League of Nations. As seen earlier, these apparent options were based on misapprehension and were not realistic possibilities in 1922. These considerations ensured that Gavan Duffy continued to champion positions that his Cabinet colleagues felt obliged to abandon. The result was a permanent sundering of relations with former political allies and a swift conclusion to his political career. Although Gavan Duffy's subsequent careers as a barrister and judge were far more successful, it is possible that the disputes that characterised his brief political career may have been responsible for a

198 UCD Archives, Gavan Duffy Papers, P152/213, memorandum (n 113 above) and *Irish Independent* (n 60 above).

199 *Dáil Debates*, vol 1, col 1461, 11 October 1922.

200 UCD Archives, Gavan Duffy Papers, P152/202, Gavan Duffy to Collins, 26 May 1922.

reputation for impracticality and pedantry within some sections of the legal professions.[201] Gavan Duffy's decision in his last days to devote the entirety of a recording of his memories to the politics surrounding the drafting of the 1922 Constitution, which by this stage was long obsolete, suggests that these events preoccupied him until the end. Gavan Duffy admitted that these memories remained painful almost three decades later.[202]

Yet, Gavan Duffy's interventions in the Constituent Assembly did have some significant consequences. It is important to remember that not all of Gavan Duffy's proposed constitutional amendments were rejected and some continue to exert influence in Irish constitutional law.[203] It is also possible that Gavan Duffy's involvement in constitutional affairs after 1932, culminating in his involvement in the drafting of the 1937 Constitution of Ireland, may have been partly inspired by his failures in the Constituent Assembly. The new Constitution was uncompromising on principles of popular sovereignty, legislative autonomy and extraterritorial jurisdiction that Gavan Duffy had championed in 1922. It also provided the opportunity for the insertion of declarations of natural law, which Gavan Duffy strongly supported, that distinguished the new Constitution from its predecessor.[204] Gavan Duffy proved so supportive of the 1937 Constitution that he would write a citizen's guide to it.[205]

Gavan Duffy was a strong advocate of cooperation with the overseas Dominions at Imperial Conferences in advancing the autonomy of the Irish Free State. This proved to be one of his most insightful contributions after the signing of the Treaty. As seen earlier, this route to constitutional reform was not nearly as obvious to contemporaries as it would be to subsequent commentators enjoying the benefit of hindsight.[206] It was openly rejected by many opponents and supporters of the 1921 Treaty who could not be sure that the restless Irish Free State would find any friends among the overseas Dominions. Gavan Duffy's argument was not helped by serious mistakes and misconceptions, in particular his belief that an Imperial Conference

---

201 Connolly (n 5 above) 129–130, 134 at 134. Connolly adds that these allegations 'must be taken with a grain of salt' and concludes that they were substantially rooted in 'earlier political antagonism stemming from the Civil War'.

202 See Military Archives (n 7 above).

203 For example, see n 163 above.

204 See n 41 above.

205 Anon [George Gavan Duffy], *Éire – The New Irish Constitution – The Citizen's Manual* (James Duffy 1938).

206 Gavan Duffy's son Colum Gavan Duffy would later argue that his father's legal arguments before the constituent assembly 'were made before their time'. C GavanDuffy (n 5 above) 95 and reprint 10.

dedicated to constitutional reform was due to be convened in the near future. Yet, although Gavan Duffy was clearly mistaken in terms of details, his instinct as to the direction of reform and the prospect of finding common cause with other Dominions proved to be entirely accurate. The Imperial Conferences of the late 1920s and early 1930s did provide the basis for sweeping constitutional change in the form of the Statute of Westminster. This facilitated the dismantling of the Treaty settlement over the course of the 1930s. Gavan Duffy's instincts on the potential offered by this peaceful avenue for constitutional change would prove to be justified, although even he may have been surprised that the settlement he had signed in 1921 would prove to be so transient.

Northern Ireland
Legal Quarterly
Summer Vol. 74.2 (2023) 362–387
Article DOI: 10.53386/nilq.v74i2.1099

Disability and COVID-19: improving legal and policy responses through grassroots disability ethics

Ivanka Antova[1]

Research Officer, Human Rights Consortium, Belfast
Correspondence email: ivanka.antova@gmail.com

ABSTRACT

The emergency legal and policy responses to COVID-19 attempt to avoid discrimination against disabled people. But they do not address deeper ableist and disableist narratives and practices embedded in emergency health policy. Adopting a disability ethics approach to the guidelines that emerged during the COVID-19 pandemic shows that they rest on dubious ethical grounds. However, emergency legal and policy responses to COVID-19 can be improved by adopting an approach based on disability ethics principles that emerge from grassroots level.

Key words: disability; COVID-19; health policy; disableism; ableism; disability ethics.

INTRODUCTION

As national health systems across the world scrambled to address the strain that the COVID-19 pandemic was expected to place on their services, especially on intensive care units (ICUs), a myriad of guidelines emerged, aiming to help medical professionals to make difficult decisions about fair and equitable distribution of scarce healthcare resources. In the United Kingdom (UK), two such key instruments are the National Institute for Health and Care Excellence (NICE) rapid COVID-19 guidelines on critical care[2] and the British Medical Association (BMA) 'COVID-19: ethical issues' guidance.[3] Both guidelines state explicitly that direct discrimination against protected categories of patients, such as elderly patients and disabled patients, is illegal, unethical and should be avoided.

1 The support of the ESRC Health Governance After Brexit project, ES/S00730X/1, is gratefully acknowledged. In particular, Professor Tamara Hervey's generous help with the structure of this article is gratefully acknowledged and appreciated.

2 NICE, 'COVID-19 rapid guidelines: critical care in adults' (NICE, 20 March 2020).

3 BMA, 'COVID-19: ethical issues' (BMA, 7 September 2020).

Yet, because they permit 'proportionate means for achieving a legitimate aim' and recognise a commitment to saving as many lives as possible as a 'legitimate aim',[4] the guidelines have the effect of leaving space for *indirect* discrimination against disabled people. As a result, the ethical guidelines have not enjoyed universal acceptance, with the disability community in particular reacting with anger to what they perceived to be 'terrifying and discriminating' guidelines.[5] These concerns remained even after the NICE guidelines were amended, in response to the threat of judicial review,[6] to provide that direct discrimination on the basis of disability is inconsistent with the legal duty of equal treatment of all patients, and that medical professionals should conduct an individual assessment of disabled patients, rather than using a 'frailty assessment method'.[7]

The anger, distrust and fear of some in the disability community may seem unfounded and misplaced. After all, the legal principles of equality and non-discrimination apply irrespective of the COVID-19 pandemic, and disabled people can rely on those principles in this time as at all times. Protection from both direct and indirect discrimination is guaranteed by the provisions of the Equality Act 2010 and by the United Nations (UN) Convention on the Rights of People with Disabilities (UNCRPD) which the UK has ratified.[8] This article challenges this perspective. Using 'grassroots disability ethics' (GDE), this article shows how ethical guidelines like the NICE and BMA guidelines embed deeper problems than those that can be resolved by a liberal equality perspective, even including indirect discrimination. GDE is understood in this article as conceptualisations and formulations of an ethical approach to emergency triage and the distribution of limited resources during the pandemic that are produced by disabled people themselves and by their organisations. GDE principles are informed by lived experiences of disability and are positioned here within the broader concept of disability inclusivity in

4 Ibid 7.

5 John Ping, 'Coronavirus: anger over "terrifying and discriminating" intensive care guidelines' (*Disability News Service*, 26 March 2020).

6 The proposed judicial review, arguing that the 'frailty' assessment method in the guidelines was an unlawful limitation on the chances of a disabled person being admitted to an ICU, was brought on the grounds of unlawful discrimination in access to critical care, quoting arts 2, 3, 8 and 14 of the European Convention on Human Rights and ss 19 and 29 of the Equality Act 2010. See Local Government Lawyer, 'NICE amends Covid-19 critical care guideline after judicial review threat' (*LGL*, 1 April 2020).

7 NICE, 'NICE updates rapid COVID-19 guideline on critical care' (NICE, 25 March 2020).

8 Equality Act 2010, art 13 (direct discrimination) and art 19 (indirect discrimination). UNCRPD, art 5 (equality and non-discrimination).

health policy. GDE principles embody a human rights-based approach to public health and health services, on account of explicitly referring to the human rights of disabled people.[9] The article then demonstrates how the principles of GDE can be harnessed for better law- and policy-making. Thus, the article develops a rationale for both critiquing and improving current ethical guidelines.

The article proceeds as follows. After a brief discussion of the broader contexts, a conceptual framework for GDE is presented. Next, the detail of the guidelines is set out. While not formally law, guidelines like these have a quasi-legal status, in that, for instance, failure to adhere to them may result in disciplinary action or a tortious claim for damages should harm to a patient ensue. The main body of the article falls into two sections. First, it analyses the guidelines using the GDE framework, explaining their deficiencies from that perspective. Second, it shows how GDE principles may provide an alternative foundation for more inclusive healthcare decision-making, in the context not only of the UK's guidelines, but also similar guidelines elsewhere, and not only of COVID-19, but also of other health emergencies and situations of scarcity in healthcare resources.

CONTEXTS: THE COVID-19 PANDEMIC AND DISABILITY

Despite the popular opinion that COVID-19 is an equalising experience that affects everyone in the same way, 'we are not all equally in this together'.[10] Disabled people are disproportionately negatively affected by the global pandemic. At international level, a Global Monitoring Report, *Disability Rights during the Pandemic*, produced by a consortium of disability rights organisations, outlines the 'catastrophic' impact of the COVID-19 pandemic on persons with disabilities. The report describes the overwhelming failures of states to take sufficient measures when responding to the pandemic to protect the rights of persons with disabilities.[11] The UN and the World Health Organization (WHO) have warned that disabled people are more at risk of contracting the virus and that some of the practical measures to stop the spread of the virus may not be possible for disabled people

9 Amanda Roberts et al, 'Treat me right, treat me equal: using national policy and legislation to create positive changes in local health services for people with intellectual disabilities' (2012) 26 Journal of Applied Research in Intellectual Disabilities 14–25, 16.

10 Katherine Hall et al, 'Ethics and equity in the time of coronavirus' (2020) 12(2) Journal of Primary Health Care 102.

11 Ciara Brennan et al, Disability Rights during the Pandemic: A Global Report on Findings of the COVID-19 Disability Rights Monitor (Global Monitoring Disability Report, 27 October 2020).

to deploy.[12] In addition, experiences of disabled people during the pandemic suggest that discrimination in critical triage is not a worst-case scenario for disabled people around the world, but a lived reality for many.[13]

At European level, the European Union (EU) Agency for Fundamental Rights' report, *COVID-19 Pandemic in the EU*, argues that the challenges that disabled people continue to face in their everyday life could even amount to discrimination in the context of the pandemic.[14] For example, complex visiting guidelines or disproportionately implemented restrictions lead to greater stress and loneliness for disabled people.[15]

More broadly, the impact of the pandemic on disabled people, and the discrimination that stems from it, should be seen in the context of historical barriers to healthcare and social care that disabled people have faced and continue to face.[16] These barriers can be physical and social (such as inaccessible buildings and inaccessible transport), communications barriers (such as lack of assistive technology), or barriers emerging from stigma and discrimination at both individual and institutional level.[17] Research on inequalities in health and social care reveals that many disabled people are discriminated against in relation to healthcare and that, despite some improvements of law and policy in the area, more progress must be made to ensure equal access to health and social care.[18]

In the UK, the disability community has expressed grave concerns about the emergency legal and policy response to the pandemic and the way that it encroaches on established disability rights. The UK pandemic response has been described as not thought-through, not

12 UN News, 'Preventing discrimination against people with disabilities in COVID-19 response' (*UN News*, 19 March 2020) .

13 Brennan et al (n 11) 43.

14 EU Agency For Fundamental Rights, COVID-19 Pandemic in the EU: Bulletin 4 (Publications Office of the EU, July 2020).

15 Ibid.

16 Ruel Serrano, 'Working to remove barriers to health care for people with disabilities' (WHO, 10 December 2012).

17 UN, *Report on the World Situation 2018*, 'Persons with disabilities: breaking down barriers' (UN Publications, 22 July 2018) ch 5.

18 See Afia Ali et al, 'Discrimination and other barriers to accessing healthcare: perspectives of patients with mild and moderate intellectual disabilities' (2013) 8(8) PLOS One; Heather de Vries McClintock et al, 'Health experiences and perceptions among people with and without disabilities' (2016) 9 (1) Disability Health Journal 74; Michael Stilman et al, 'Healthcare utilization and associated barriers experienced by wheelchair users: a pilot study' (2017) 10(4) Disability and Health Journal 502; Dora Raymarker et al, 'Barriers to healthcare: instrument development and comparison between autistic adults and adults with or without other disabilities' (2017) 21(8) Autism 972.

proportionate and not protecting everyone.[19] The Coronavirus Act 2020,[20] the overarching legal response to the pandemic, is deeply problematic from a disability perspective, in particular because it removes the statutory duty on local authorities to provide social care services during the pandemic.[21] Writing in 2013, the prominent disability scholar and activist Mike Oliver can now be seen as prophetic about the way in which the Coronavirus Act 2020, as a response to COVID-19, has changed disability rights:

> Our differences are being used to slash our services as our needs are now being assessed as being moderate, substantial or critical and many local authorities are now only providing services to those whose needs are critical.[22]

These are the contexts in which guidelines for medical professional practice in the context of COVID-19 were developed.

GRASSROOTS DISABILITY ETHICS

The historic barriers to healthcare outlined above, as well as the challenges brought by the COVID-19 pandemic, have strengthened the call for a disability-inclusive approach to the ongoing public health crisis. GDE principles for an ethical distribution of limited resources fit within the broader concept of a disability-inclusive approach to health and health policy. A disability-inclusive approach is understood as 'mainstreaming disability in all plans and efforts', as well as 'adopting targeted measures' that meet specific requirements, since general responses to the pandemic might not respond effectively to the particular needs of disabled people.[23] Such an approach is inherently person-centred. It calls for the effective inclusion of disabled people as active participants in deciding how to meet their needs during the pandemic, alongside a core group of stakeholders, including family members and health professionals.[24] Disability inclusivity in emergency responses is facilitated by challenging the 'morally reprehensible' deprioritising of disabled people during the pandemic with a strong focus on their

19 John Pring, ' Coronavirus: grave concern over impact of emergency Bill on rights' (*Disability News Service*, 19 March 2020).

20 Coronavirus Act 2020.

21 Ivanka Antova, 'Disability and COVID-19 in England: emergency policy and legal responses' (2020) 28(4) Medical Law Review 804–816.

22 Mike Oliver, 'The social model of disability: 30 years on' (2013) 28(7) Disability and Society 1024, 1026.

23 UN, Policy Brief: A Disability-inclusive Response to COVID-19 (2020) 8.

24 S Senjam, 'A persons-centered approach for prevention of COVID-19 disease and its impacts on persons with disabilities' (2021) 8 Frontiers in Public Health 3.

human rights.[25] As the pandemic lays bare the disproportionately negative impact of COVID-19 on disabled people, disability scholars have renewed the pre-pandemic call for radical changes to be made to the way disability policy and health policy are made by means of mainstreaming disability lived experiences.[26]

The inclusion of disabled people's voices and lived experiences in 'both direct and indirect measures in the fight against COVID-19' is a legal requirement.[27] Article 4(3) of the UNCRPD requires states to 'closely consult and actively involve' disabled people and their representative organisations in the implementation of the Convention. In addition, article 33(3) requests states to ensure that disabled people participate fully in the monitoring of the implementation of the Convention. Disabled people and their organisations have been involved in the very creation of the UNCRPD through a 'unique' approach to treaty drafting that affords equal status to civil society members and state representatives, thus giving the UNCRPD an 'edge it would otherwise have lacked' had it not incorporated the lived experience of disability.[28] The effective inclusion of disability in the creation of international and domestic human rights standards and health protocols has been seen as a key step towards 'reshaping present exclusionary practices and structures' that underpin to a large extent the disproportionately negative impact of the COVID-19 pandemic on disabled people.[29]

There are real practical merits in ensuring the effective participation of disabled people in the drafting of emergency responses to the pandemic. The lived experience of disability can inform practices that mitigate some of the negative impact of the pandemic by ensuring that emergency measures are 'appropriately tailored' for disabled

25 Hannah Kuper, Lena Morgon Banks, Tess Bright, Calum Davey and Tom Shakespeare, 'Disability-inclusive COVID-19 response: what it is, why it is important and what we can learn from the United Kingdom's response' [version 1; peer review: 2 approved] (2020) 5:79 Wellcome Open Research 3.

26 Laufey Löve, Rannveig Traustadótti, Gerard Quinn and James Rice, 'The inclusion of the lived experience of disability in policymaking' (2017) 6–33 Laws 1–16, 2.

27 Elena S Rotarou, Dikaios Sakellariou, Emily J Kakoullis and Narelle Warren, 'Disabled people in the time of COVID-19: identifying needs, promoting inclusivity' 11:03007 (2021) Journal of Global Health, 3

28 Löve et al (n 26) 3.

29 Ieva Eskyte, Anna Lawson, Maria Orchard and Elizabeth Andrews, 'Out on the streets – crisis, opportunity and disabled people in the era of Covid-19: reflections from the UK' (2020) 14 European Journal of Disability Research 329–336, 334.

people.[30] There have been some attempts to do this in the UK during the pandemic. Examples include the provision of information about COVID-19 and social-distancing measures in British Sign Language and Easy Read formats for those with intellectual impairments;[31] or the development of guidance for carers of people with specific conditions, such as Alzheimer's.[32] The effective inclusion of disabled people's voices will also be important for stepping into the 'new normal' of a post-pandemic world where the less-known long-term effects of COVID-19 are studied and addressed.

Although GDE principles for the ethical and non-discriminatory response to the pandemic fit within the broader frameworks of disability inclusivity, person-centred and human rights-based approaches to health policy, GDE principles are understood in this article as potentially further reaching. GDE calls for the elevation of grassroots disability narratives to the main source, or foundation, of the emergency legal and policy responses. In this sense, a more appropriate conceptual framework to highlight the potential of GDE to achieve the 'reform in both the process and direction of policymaking' that the COVID-19 pandemic necessitates is the concept of co-production.[33]

Co-production, or 'the involvement of patients, service users, and members of the public in the design and delivery of healthcare' is an example of a grassroots disability activist narrative that has been gradually mainstreamed in policymaking.[34] Co-production goes beyond the call for effective inclusion of disabled people in health policy drafting and focuses on reversing the power disbalance within disability policy by placing disability lived experience as the leading expertise. A key element of GDE principles as co-production of emergency responses to COVID-19 (and to any future crisis) is the transformation of disabled people from passive recipients of legal and policy responses to active participants in 'collective organisational co-management and co-

30 Lieketseng Ned, Emma Louise McKinney, Vic McKinney and Leslie Swartz, 'COVID-19 pandemic and disability: essential considerations' (2020) 18(2) Social and Health Sciences 143.

31 MENCAP has produced an Easy Read summary of the government COVID-19 guidance from May 2021: The Coronavirus Rules from Monday 17th May.

32 The Alzheimer's Society has produced advice and guidance specifically for carers of people with dementia: 'Helping a person with dementia to keep safe and well during coronavirus'.

33 Peter Beresford, 'What are we clapping for? Sending people to die in social care: why the NHS did this and what needs to happen next?' in Peter Beresford et al (eds), *COVID-19 and Co-production in Health and Social Care Research, Policy, and Practice Volume 1: The Challenges and Necessity of Co-production* (Polity Press 2020) 94.

34 Nicola Gale, Patrick Brown and Manbinder Sidhu, 'Co-production in the epidemiological clinic: a decentred analysis of the tensions in community-based, client-facing risk work' (2018) 53 Social Policy Administration 203–218, 204.

governance of health'.[35] This is crucially important for challenging definitions of disability as individual physiological failure, as worthlessness or as a societal burden. GDE principles thus have the capacity to radically disrupt 'wider social and cultural processes that disempower and exclude' in favour of the sharing of decision-making with the community that empowers individuals and protects human rights.[36] During the pandemic, disabled people experience more than a higher risk of exposure to COVID-19, or lockdowns and social-distancing measures incompatible with their lives. Disabled people experience emergency responses to the pandemic that allow for indirect discrimination and inherently disempower and exclude them. The co-production of emergency legal and policy responses would allow for lived experience of disability to illuminate potentials for discrimination and produce truly effective protocols. Therefore, a GDE approach to the ethical guidelines that emerged during the pandemic is a key tool in both critiquing and improving legal and policy responses. The effectiveness of GDE in critiquing emergency guidelines is discussed next.

THE GUIDELINES

Both the NICE rapid COVID-19 guidelines on critical care[37] and the BMA's 'COVID-19: Ethical Issues' guidance[38] envisage difficult choices about prioritising patients having to be made by medical professionals only in a situation where the health system, or a particular hospital or ICU within it, is overwhelmed. In such a scenario, as the guidelines point out, however undesirable this might be, prioritisation of patients will become inevitable.

The BMA guidelines explicitly commit to each patient receiving the highest possible level of care during the pandemic. The BMA guidelines go on to balance two different, and competing, approaches to the distribution of limited resources. On the one hand, there is respect for the individual and the individual right to health. On the other hand, there is a utilitarian concern for the health of the population as a whole. If sufficient resources become unavailable, then utilitarianism must prevail, and the leading concern should be to minimise overall mortality and morbidity.

35 Andrew G H Thompson, 'Contextualising co-production and co-governance in the Scottish National Health Service' (2020) 5(1) Journal of Chinese Governance 48–67, 49.

36 Jane Booth, 'Empowering disadvantaged communities in the UK: missing the potential of co-production' (2019) 49 (2) Social Change 276–292, 282.

37 NICE (n 2).

38 BMA (n 3).

> Although doctors would find these decisions difficult, if there is radically reduced capacity to meet all serious health needs, it is both lawful and ethical for a doctor, following appropriate prioritisation policies, to refuse someone potentially life-saving treatment where someone else is expected to benefit more from the available treatment.[39]

The guidelines acknowledge that, whilst this situation would necessitate difficult and possibly distressing decisions, age and disability on their own may not be the only factors to be taken in consideration. Decisions should be based on 'evidence and reason'. However, in some cases, age and disability may feature as part of such an evidenced and reasoned decision-making process.

What medical professionals should prioritise in these very challenging circumstances is a higher survival probability, and a consideration of which patients would be expected to benefit more from critical care. The most urgent cases, the least complex cases, and patients expected to live the longest as a result of receiving critical care should be prioritised. Patients with co-morbidities that would impact on their capacity to benefit from treatment should not be prioritised. Patients who have 'sufficient background illness' or those who are frail should not be prioritised. Long-term health conditions are seen as a reason not to prioritise, while the key factor for prioritisation should be the capacity to benefit quickly from treatment.[40]

Although the guidelines recognise the key principle of reasonable adjustment as an important part of disabled people's equal access to health care, they envisage a scenario where this duty is affected by the pandemic. The guidelines' position is that reasonable adjustment should not 'trump' the utilitarian commitment to saving as many lives as possible. To this end, and only in this limited context, indirect discrimination, or unintentional discrimination against disabled patients because of their difference from other patients, would be ethical and lawful medical practice.

The NICE guidelines make similar arguments, but in a more broad-brush way. The NICE guidelines do not provide a detailed explanation of how decisions about whom to prioritise should be made in a situation where resources are insufficient. Instead, the NICE guidelines focus more heavily on the clinical factors that should be prioritised in decision-making.

The NICE guidelines state that, when making a decision about admitting a disabled patient to critical care, medical professionals should do two things. First, they should conduct an individual

39 Ibid 3.
40 Ibid.

assessment of frailty.[41] According to the NICE website, frailty is described as 'a loss of resilience that means people don't bounce back quickly after a physical or mental illness, an accident or other stressful event'.[42] This definition is based on a British Geriatrics Society model for recognising and assessing frailty.[43] Since the NICE guidelines are for admitting patients into ICUs, medical professionals would be able to make such assessments, and this suggests that the critical frailty score might still be used when deciding which patient should be prioritised in a situation of limited health resources.

The second thing medical professionals should do is follow the algorithm that the guidelines provide. According to the algorithm, if a patient is considered to be less frail, by using the individualised assessment described above, then admission into critical care is seen as appropriate. If the patient is deemed to be more frail, then a doctor must make an additional decision about admission to critical care as part of a holistic assessment. Although we have no detailed description of what a holistic assessment might mean, we can see from the guidelines that medical professionals should always consider co-morbidities, underlying health conditions, pathologies and severity of acute illness when deciding whom to prioritise.[44]

ANALYSIS

The guidelines do not suggest that disabled patients should automatically be excluded from receiving critical care, nor do they make an explicit argument that disabled lives do not matter. But, although they proclaim that discrimination against disabled people is not permitted, the BMA and NICE guidelines nonetheless embody and articulate a highly problematic approach when seen from the point of view of GDE. This is the case for five main overlapping reasons:

- the approach of the guidelines to the balance between utility and equality;
- the construction of disability as abnormality;
- disability as representing low quality of life or health;
- a concept of the 'ideal patient'; and

41 NICE (n 2) (my emphasis). Note that the algorithm states that 'any patient aged under 65, or patient of any age with stable long-term disabilities (for example, cerebral palsy), learning disabilities or autism: do an individualised assessment of frailty. Do not use CFS score.'

42 NICE, 'Improving care and support for people with frailty: how NICE can support local priorities'.

43 Jill Turner, 'Recognising frailty: good practice guide' (British Geriatrics Society, 11 June 2014).

44 NICE (n 2).

- the way the guidelines construct a disableist response to the COVID-19 pandemic from an ableist perspective.

Each is now discussed in turn.

The utility/equality balance

In the context of a global pandemic, or other health emergency, the necessity to prioritise limited resources inherently dictates that judgements about the value of lives will have to be made. To this end, the guidelines embody what has been described as an 'unstable compromise'[45] between competing ethical approaches, namely utilitarianism and egalitarianism. The scenario where the overall health of the population is seen as competing with the health of individuals is a fertile ground for negative conceptions of disability as a 'product of a damaged body or mind'[46] to underpin decisions about who should be saved and who should not be.

The guidelines state that utilitarian concerns override commitment to prioritising each patient, regardless of how their individual health might be perceived or valued. When developing an ethical reasoning or practice for distributing limited resources, the guidelines adopt an approach based on orienting activity toward a utilitarian good. In effect, this utilitarian good amounts to a devaluing of disabled lives, as less worthy of public investment.[47] Consistently with public health ethics, the utilitarian approach typically prioritises young and healthy people.[48] The BMA guidelines in particular make an explicit call for the overall morbidity and mortality being minimised, by allowing for disability to feature as a decision-making factor when choosing the patients in whom limited resources should be invested.

Such a utilitarian approach prioritises 'normal' and 'healthy' lives, and the overall health of a nation. A categorical exclusion, understood as a manifestation of the utilitarian principle of maximising population outcomes, would exclude patients with certain co-morbidities (for instance, severe cognitive impairment) as a priority for critical care.[49]

---

45 Julian Savulescu, James Cameron and Dominic Wilkinson, 'Equality or utility? Ethics and law of rationing ventilators' (2020) 125(1) British Journal of Anaesthesia 10.

46 Dan Goodley and Katherine Runswick-Cole, 'The violence of disablism' (2011) 33(4) Sociology of Health and Illness 602, 603.

47 Shane Neilson, 'Why I won't see you on the barricades' (2020) 66 Canadian Family Physician 448, 450.

48 Jerome Singh and Keymanthri Moodley, 'Critical care triaging in the shadow of COVID-19: ethics considerations' (2020) 110(5) South African Medical Journal 355, 355.

49 Douglas White and Bernard Lo, 'A framework for rationing ventilators and critical care beds during the COVID-19 pandemic' (2020) 323(18) Journal of the American Medical Association 1773, 1773.

It would be more ethical to prioritise a non-disabled person because they would be understood as able to make a better contribution to the overall health of society, after recovery. But this population-focused utilitarian approach runs the risk of turning 'critical care into a life raft: the vulnerable are thrown overboard to keep the ship afloat'.[50]

An alternative, less restrictive, non-categorical utilitarian approach would focus on universal eligibility for critical care, but would see a prioritisation based on who would be 'most likely to benefit'. The people most likely to benefit are seen as those who would survive to hospital discharge if given the treatment. For example, it would be ethical to prioritise those with more years left to live, whether disabled or not. Or it would be ethical to prioritise younger patients, whether disabled or not, in order to give everyone an equal chance of going through all life stages (the life-cycle principle).[51] A GDE approach to the guidelines reveals that even this utilitarian approach runs the risk of perceiving disabled people as less likely to benefit from treatment because of how disability is understood, as opposed to an able-bodied or cognitively able 'norm'.

Disability as abnormality

When disability is seen as failure or abnormality, the life of a disabled patient is unlikely to be valued as much as a life that is considered 'normal' and a part of the health of a 'normal' society. Grassroots disability narratives have long challenged the portrayal of disability as abnormality and as an individual tragedy, rather than the end result of structural barriers and inequalities. From a GDE perspective, the COVID-19 ethical guidelines can be understood as a continuation of the long-standing discussion within disability studies about the prevalence of medical conceptions of disability, as opposed to conceptions of disability that aim to distance disability from biological determinism and functionalism (broadly speaking the social model of disability).

The medical model of disability, also referred to as the individual or 'personal tragedy' model of disability, is an early theory of disability that emerged from the medical profession, with the medical knowledge on the functions or performances of the body constructing disability

50 Andrew Peterson, Emily Largent and Jason Karlawish, 'Ethics of reallocating ventilators in the Covid-19 pandemic' (2020) 369 British Medical Journal 1, 1.

51 White and Lo (n 49).

as a failed performance or abnormally functioning body or mind.[52] Goodley describes what is understood here as a medical model as the dominance of functionalism as a social theory, which sees disability as the product of a damaged body or mind, thus 'functionalism underpins ableism: the social, cultural and political conditions of contemporary life that emphasise ability and denigrate disability'.[53] The medical model works through individualisation of disability.[54] The human body is seen as a 'universe' in itself and the 'problems' of this body are limited by the physical contours of the body, not to be equated with problems that a population or a group of people might experience collectively. Therefore, the medical model of disability is a highly divisive way of thinking: 'within the purview of the medical establishment, to keep it a personal matter and "treat" the condition and the person with the condition rather than "treating" the social processes and policies that constrict disabled people's lives'.[55] The COVID-19 pandemic presents significant challenges to disabled people that necessitate a deeper understanding of how disabled lives should be protected than the medical model affords.

By contrast, the social model of disability, which has become the normative analytical framework for disability studies, separates impairment from disability and places disability as the end result of the barriers that society creates.[56] The social model emerged as a framework to make sense of disability in 1976 in the work of the Union of the Physically Impaired Against Segregation (UPIAS), a group of disabled activists and socialists. Mike Oliver further elaborated the

52 Marno Retief and Rantoa Letšosa, 'Models of disability: a brief overview' (2018) 74(1) HTS Teologiese Studies/Theological Studies a4738, 3. See also Andrew J Hogan, 'Social and medical models of disability and mental health: evolution and renewal' (2019) 191(1) Canadian Medical Association Journal E16–E18; Jonathan M Levitt, 'Developing a model of disability that focuses on the actions of disabled people' (2017) 32(5) Disability and Society 735–747; and Stephen Bunbury, 'Unconscious bias and the medical model: how the social model may hold the key to transformative thinking about disability discrimination ' (2019) 19(1) International Journal of Discrimination and the Law 26–47.

53 Goodley and Runswick-Cole (n 46) 603.

54 Joel Michael Reynolds, '"I'd rather be dead than disabled" – the ableist conflation and the meanings of disability' (2017) 17(3) Review of Communication 149–163, 151.

55 Simi Linton, *Claiming Disability: Knowledges and Identity* (New York University Press 1998) 4.

56 Jonathan Levitt, 'Exploring how the social model of disability can be re-invigorated: in response to Mike Oliver' (2017) 32(4) Disability and Society 589.

social model in 1983[57] and 1990,[58] later describing it as nothing more than 'a tool to improve people's lives'.[59]

In the UK, the social model of disability has indeed been a highly effective tool for disability activism and emancipatory disability narratives (within which GDE can be placed), despite critiques of the social model and many of its limitations having been discussed at length.[60] Perhaps most importantly, critical disability studies scholars looking to go beyond the social model have argued that 'bodies are not simply born, but made'.[61] The strict separation of impairment from disability could leave the disabled body (or mind) open to theoretical interventions and definitions from a medical perspective alone, rendering disability a personal tragedy or failure, rather than an experience affecting many.[62] A GDE approach would instead prioritise disability-led, inclusive and human rights-centred definitions of disability, in line with the social model of disability.

The BMA and NICE guidelines can be seen as emergency responses that have inherited the medicalisation of disability, which dominates the medical profession. By relying heavily on utilitarian principles to justify indirect discrimination against disabled patients, the ethical guidelines in effect prioritise 'normality' when decisions about who should receive scarce healthcare resources are made. Thus, the guidelines reinforce the notion that 'abnormality' can and should be excluded if resources are limited. As such they are an embodiment of the medical model of disability that many disabled people see as undermining the validity of their existence.

The WHO requires states to 'ensure that decisions on the allocation of scarce resources (eg ventilators) are not based on pre-existing impairments, high support needs, quality of life assessments, or medical bias against people with disability'.[63] But guidelines like the BMA and NICE guidelines do not provide the necessary clarity[64] to medical professionals, especially where they lack knowledge and have

57 Mike Oliver, *Social Work with Disabled People* (Macmillan 1983).

58 Mike Oliver, *The Politics of Disablement* (Macmillan 1990).

59 Oliver (n 22) 1025.

60 Janine Owens, 'Exploring the critiques of the social model of disability: the transformative possibility of Arendt's notion of power' (2014) 37(3) Sociology of Health and Illness 385.

61 Elizabeth Donaldson, 'The corpus of the madwoman: toward a feminist disability studies theory of embodiment and mental illness' (2002) 14(3) Feminist Disability Studies 99, 112.

62 Ibid.

63 WHO, Disability Considerations during the COVID-19 Outbreak (WHO 2020).

64 Richard Huxtable, 'Bin it or pin it? Which professional ethical guidance on managing COVID-19 should I follow?' (2020) 21(1) BMC Medical Ethics 1, 9.

insufficient training in the needs and rights of disabled people.[65] In those circumstances, where medical professionals conduct the individual assessment required by the guidelines when deciding whom to prioritise for limited resources during the pandemic, such an assessment would involve an individual disabled person being evaluated and labelled through a process which separates that disabled person from mainstream society, education, work or social interaction. The separation arises because a disabled person is seen as deviating from an implicit dominant norm, and their difference is not valued.[66]

Disability as low quality of life or health

In a similar way, the guidelines embody the idea that disability is associated with a low quality of life, or health, when compared to an able 'norm'. Within the medical model, disability is understood as inherently negative, something to be endured, which should be cured or even eliminated, if possible.[67] A GDE perspective would offer an understanding of life with a disability as something that might be experienced, or even enjoyed, as a normal part of the life of an individual. When disabled people enter the medical field, they encounter difficulties or barriers because quotidian experiences for them (such as the use of feeding tubes or respirators) become indicators of an unacceptably low quality of life.[68] Including quality of life as a factor 'risks incorporating concerning value judgments that will systemically disadvantage people with disabilities and chronic health conditions and reduce the likelihood that they will receive medically indicated care'.[69] The perception of disability as an inevitable prognosis for bad quality of life post-critical care, or as a negative prognosis in terms of a fast recovery, allows for disabled patients to be deprioritised for access to a ventilator, even if they need it more than a non-disabled person presenting with the same disease.[70]

As the guidelines are applied by medical professionals, disabled people's impairments, or underlying health conditions that may be the reason for their disability, will be seen as a medically relevant ground for exclusion from prioritisation of resources. This is the case

65 Maya Sabatello et al, 'People with disabilities in COVID-19: fixing our priorities' (2020) 20(7) American Journal of Bioethics 187, 187.

66 Owens (n 60).

67 Fiona Campbell, *Contours of Ableism: The Production of Disability and Abledness* (Palgrave MacMillan 2009) 5.

68 Heidi Jenz, 'Ableism: the undiagnosed malady afflicting medicine' (2019) 191 Canadian Medical Association Journal E478, E479.

69 Ari Ne'eman, 'When it comes to rationing, disability rights law prohibits more than prejudice' (The Hastings Centre, 10 April 2020) 2 .

70 Gerard Goggin and Katie Ellis, 'Disability, communication, and life itself in the COVID-19 pandemic' (2020) 29(2) Health Sociology Review 168, 171.

even though disabled patients are seeking ICU admission because of COVID-19 and not because of a stable or a long-term health condition that they would otherwise manage. Disability, when understood only as reduced capacity to function 'normally', may be seen as inherent frailty, or as incapacity to be 'healthy' despite access to critical care.

The merit of applying a geriatric model of frailty assessment to disabled patients is highly questionable, since old age and disability may overlap in certain cases, but are certainly not the same thing. Frailty is assessed by various means, for example looking into the speed of walking, the strength of a grip, any increased challenges in getting out of bed or going to the toilet. For some disabled patients with particular health conditions or impairments these challenges may be an everyday reality, not necessarily a signal of increased frailty that may be seen as a reason not to be prioritised for critical care during a pandemic. Even the British Geriatrics Society model allows for the use of the critical frailty score, *only after* a comprehensive individual clinical assessment.

This concern about a frailty model lies at the heart of the successful challenge to the original NICE guidelines, which were amended precisely because they equated disability with frailty and frail patients were to be excluded from receiving critical care in favour of less frail, or non-disabled patients. The now amended NICE guidelines call on medical professionals to recognise the limitations of assessing disability as frailty and insist on an individualised assessment to be carried out instead. But it is precisely during this individualised assessment that the perception of disability as low quality life or health enters the decision-making process. Indirect discrimination against disabled people takes place when disabled lives and experiences are measured against an unattainable 'ideal' notion of health and humanity that distinguishes between disabled people and non-disabled people and leaves the former in a disadvantageous position. The ethical guidelines require attention to quality of life post-treatment. A GDE approach to the guidelines reveals that, where disabled lives are seen as lower quality and disabled people perceived as having a lower quality of health, the guidelines steer resourcing decisions in a way that discriminates against disabled people.

The 'ideal patient'

The NICE and BMA guidelines suggest that patients without underlying health conditions and co-morbidities, with a better ability to survive, with a better chance of benefiting from treatment and, perhaps most importantly, with less complexity to their health circumstances, would be *ideal* for prioritisation during the pandemic. In other words, the guidelines prioritise those who are seen as healthier already,

or those patients who are seen to be the closest to the unattainable health performance that constructs disability as lacking. From a GDE perspective, this is problematic. It articulates a normative notion of a rational, independent, autonomous subject, embedded in the notion of an ideal human patient. Such an 'ideal patient' is often evoked in policy making,[71] and COVID-19 guidelines are no exception. The guidelines lean heavily towards protecting this normative construction that leaves disabled patients, who inevitably challenge the hegemony of the 'ideal patient', in a disadvantageous position.

The pandemic constructed from an 'ableist' and 'disableist' perspective

Making this point more broadly, law and policy processes that construct human bodies according to an 'ideal' also contribute to ways that pandemic responses, like the BMA and NICE guidelines, embody and articulate an approach that is highly problematic from the point of view of GDE.

The ideal patient described above, the one who should be prioritised for limited health resources during a pandemic, is an example of defining disability as the opposite of ideal or normal. From a critical disability perspective, defining disability is about destabilising a definition of disability that relies on normative idealised understandings of how the human body or mind *should* perform. To build this destabilising narrative, critical disability scholars have developed the concepts of 'ableism' and 'disableism' as the two sides of the same coin, mutually supporting and promoting each other.[72] Goodley defines ableism as an ideal, not something to which anyone ever matches up. Disableism is the process of pressing normative ableism upon people: 'the oppressive practices of contemporary society that threaten to exclude, eradicate and neutralise those individuals, bodies, minds and community practices' that do not reach the unattainable ableist standard.[73] These practices occur across social contexts, including in the medical field. From this perspective, two entirely distinct categories exist: disabled or able,[74] the latter of which does not exist in absolute form, but as an imagined norm to which people can be compared. Ableist-normativity

71 Dan Goodley and Katherine Runswick-Cole, 'Becoming dishuman: thinking about the human through dis/ability' (2014) Discourse: Studies in the Cultural Politics of Education 2.

72 Dan Goodley, *Dis/ability Studies: Theorising Disablism and Ableism* (Routledge 2014) ix.

73 Ibid xi.

74 Campbell (n 67) 8.

is constantly produced and maintained through such comparisons,[75] ensuring that disableist discrimination takes place, by implicitly casting disabled people as comparatively 'less-than-human'.[76]

The COVID-19 ethical guidelines overtly favour those patients who are closer to the ableist normativity of the healthy individual, by marrying body performativity that challenges the notion of 'being healthy' with relevant medical factors that determine who should be prioritised for the limited health resources. The systemic nature of ableism and disableism is reproduced in the guidelines:

> Governmental policies, laws, and rules ... are designed for the benefits of the privileged group, people without impairment or disability. Ableism is constructed on the basis of hierarchy where people without disability are on the top.[77]

According to the guidelines, medical professionals should prioritise a higher survival probability, those expected to benefit more from treatment, the least complex cases, patients without co-morbidities or background illness, those most likely to recover, and those with capacity to recover quickly. These factors may seem perfectly relevant and objective to a medical professional, or to a lay person. The argument here is not that medical professionals would deliberately discriminate against disabled people. It is rather that medical professionals rely on an ableist definition of disability as a fixed or stable body or mind that simply does not perform as well as a non-disabled body or mind. In that context, the perception that medical intervention is less likely to 'fix' a disabled person's chances of recovery, even if critical care is administered, appears rational.

From a GDE perspective, however, when understood as incorporating the lived experience of the effects of processes of ableism/disableism, the guidelines fail to recognise or reflect the much more complex and fluid range of disability experiences. The medical profession is not exempt from producing ableist narratives and disableist practices under the disguise of medical knowledge or 'common sense'. '[A]bleism is that most insidious form of rhetoric that has become reified and so widely accepted as common sense that it denies its own rhetoricity—it "*goes without saying*".'[78]

---

75 Fiona Campbell, 'Refusing able(ness): a preliminary conversation about ableism' (2008) 11(3) Media and Culture 1.

76 Laura Sanmiquel-Molinero and Joan Pujol-Tarrés, 'Putting emotions to work: the role of affective disablism and ableism in the constitution of the dis/abled subject' (2020) 33(4) Disability and Society 602, 605.

77 Heeson Jun, *Social Justice, Multicultural Counselling, and Practice: Beyond a Conventional Approach* (Springer International 2018) 246.

78 James Cherney, 'The rhetoric of ableism' (2011) 31(3) Disability Studies Quarterly.

Medical criteria that may be accepted as common sense, or, as the guidelines embody, the implicitly inevitable processes that doctors have to follow in the context of resource shortage, are not necessarily objective or value-free. Rather, they can be understood as an example of medical ableism exerting pressures during the pandemic that allow for disableism in the form of legal and ethical discrimination against disabled people to be justified within the guidelines. Critical care could be denied a disabled person, consistently with the guidelines, based on an ableist perception of the disabled person's quality of life or health applied during a triage process: a third party (medical professional) concludes that the disabled person's life has insufficient quality to be worth saving in comparison with non-disabled (or, rather, less disabled, as 'able' is an unattainable norm) others.[79] What may be considered as 'common sense' or medical objectivity, from a GDE perspective is revealed as smuggling in judgements on quality of life. These judgements are particularly pernicious when 'health' is nebulously defined as 'well-being', in effect a synonym for quality of life.[80]

Equally, consideration of long-term survival and short-term survival as a relevant factor is problematic from a grassroots disability perspective. Decisions, at least on allocation of treatment modalities and hospital beds, based on long-term or short-term survival have been seen as appropriate in the context of COVID-19.[81] But these have the effect of discriminating against disabled people, who regularly outlive the prognosis ascribed by medical professionals, often by decades. A disability does not automatically indicate a poor prognosis for short or longer-term survival.[82]

To summarise: from the point of view of GDE, the guidelines can be understood as constructing responses to the COVID-19 pandemic from an ableist perspective, which undervalues equal treatment of different bodies (or minds) in the name of a utilitarian approach to individual and societal health, 'well being' or quality of life that is constructed by reference to an 'ideal patient' who does not embody an 'abnormal' disability. Thus, the guidelines permit processes of decision-making that have the effect (even if not the intention) of discriminating against disabled people.

---

79 Joseph Stramondo, 'COVID-19 triage and disability: what NOT to do' (*Bioethics Net*, 30 March 2020).

80 Anthony Gavin, 'The poverty of bioethics: disability, medical austerity, and traumatic care' (*Social Sciences and Humanities Open*, 4 June 2020).

81 Naomi Laventhal et al, 'The ethics of creating a resource allocation strategy during the COVID-19 pandemic' (2020) 146(1) Pediatrics 4.

82 Disability Rights Education Defense Fund (DREDF), 'Applying HHS' guidance for states and health care providers on avoiding disability-based discrimination in treatment rationing' (*Disability Rights Education and Defense Fund*, 3 April 2020).

AN ALTERNATIVE APPROACH: GDE

It is one thing to critique a law or policy from an external standpoint. It is quite another to offer an alternative approach. This article argues that GDE provides not only a standpoint for critique, as shown above, but also a basis for an alternative approach to healthcare decision-making in the context of COVID-19, and indeed in broader contexts, involving healthcare emergencies and/or scarcity of healthcare resources such as ICUs. GDE has potential to provoke positive change in three contexts: healthcare practice, healthcare policy and disability ethics more generally.

To illustrate this potential, let us first contrast the position of two doctors, from the UK and Canada, in terms of how they understand the decision-making required of them under ethical guidelines such as the BMA or NICE Guidelines.

Consider first the perspective of Dr Matt Morgan, a UK -based NHS doctor from Cardiff. Dr Morgan wrote an open letter from an ICU 'to those who are elderly, frail, vulnerable, or with serious underlying health conditions'.[83] In this remarkable letter, Dr Morgan reassures these people that they have not been forgotten. But observe the way in which the role of the medical professional is described during the pandemic:

> Our passion as an intensive care community is *fixing problems that can be fixed*. Yet we often meet patients like you who have problems that *cannot simply be fixed* … As difficult as this is, we will be honest. We will continue to use all of the treatments that may work and may get you back to being you again. We will use oxygen, fluid into your veins, antibiotics, all of the things that may work. But we won't use the things that won't work. We won't use machines that can cause harm. We won't press on your chest should your heart stop beating. Because these things won't work. They won't get you back to being you.[84]

To Dr Morgan, the point of doctors is to 'fix' patients. The challenge that disabled patients present is that they cannot be 'fixed'. In the context of a global pandemic and limited resources, this challenge becomes more acute. Not even powerful technology and advanced medication can fix the problem. The end strategy seems to be a nod towards 'do not resuscitate'.[85] Compassion, care, attention and understanding of disabled people is not denied here, quite the opposite. There is a nod to individual experience: 'you being you'. But there is a stronger sense of disableist inevitability in Dr Morgan's words that the responses to

83 Matt Morgan, 'Letter from ICU' (*The BMJ Opinion*, 12 March 2020).

84 Ibid (my emphasis).

85 John Pring, 'Coronavirus: activists' shock at intensive care doctor's resuscitation warning' (*Disability News Service*, 19 March 2020).

the pandemic will be impacting negatively on disabled people and a resignation or even acceptance that this cannot be changed, or 'fixed'.

Now, by contrast, consider the perspective, of Dr Shane Neilson, a Canadian doctor and scholar, and a father of a disabled son. In his opinion piece ('Why I won't see you at the barricades'), Dr Neilson discusses the 'vexed' relationship between the medical profession and disability, one that has historically disadvantaged and discriminated against disabled people. Dr Neilson responds to the call of duty to the profession to intensify work efforts during the pandemic and make difficult decisions on prioritising resources (to go to the barricades) with the following:

> In truth, I love to go to work, but not for you, not exactly; not for an abstract ideal; definitely not for emergency services vehicles sounding their klaxons in a fluid cordon around a building. I do it for me, because I like doing it, love it in fact. I do it because I like helping someone else; it makes me feel good. But the second my work becomes an activity oriented toward a utilitarian good, a recruited assent toward devaluing disabled lives, and a requirement I place myself at greater risk (and thereby my family, including my disabled son), I say no.[86]

Dr Neilson, in contrast to Dr Morgan, acknowledges the inherent unfairness of the necessity to prioritise resources for patients who can be 'fixed'. Dr Neilson does not accept the disableist inevitability of the medical profession having to make these decisions with a compassionate confidence in their ethical soundness. In fact, he calls for action and resistance: 'When disabled lives are explicitly protected by a discipline that historically has preferentially extinguished them – *that's* when I'll join you at the barricades.'[87]

GDE principles are principles for the ethical, fair and just distribution of limited resources that disabled people produce themselves based on their lived experience, and hoped-for futures,[88] and expertise in navigating through the complexities of ableism, disableism and different models of disability described above. Far from being passive recipients of ethically questionable emergency responses to the global pandemic, the disability community has been active in its resistance to the brutality of utilitarianism, to the reductionism inherent in bio-economic decision-making, and to the expressions of the value or worth of human lives these kinds of responses, including as embodied in the BMA and NICE guidelines, entail.[89]

---

86 Neilson (n 47) 450.

87 Ibid 450.

88 Ibid 449.

89 Thomas Abrams and David Abbott, 'Disability, deadly discourse, and collectivity amid coronavirus (COVID-19)' (2020) 22(1) Scandinavian Journal of Disability Research 168.

Perhaps the most valuable contribution GDE can make to the evolving subject of ethics during the COVID-19 pandemic is in recognising the importance and prioritising the inclusion of lived experience in legal and policy responses to this pandemic, as well as future public health crises. Guidelines for provision of healthcare during COVID-19 must be developed in collaboration with disabled people's organisations and representatives from human rights bodies. Disability ethics based on lived experience can be a valuable tool for overcoming ideological divides and ethical disagreements, especially those which are framed as in-principle zero-sum decisions. The lived experience has a transformative power in ethical contexts. Instead of talking about abstract or theoretical concepts, to be solved by logical consistent argument, the conversation becomes about a set of concrete problems to be solved with practical reform informed by real people's experiences.[90] The medical profession has historically excluded disability voices, experiences and deconstructions of normative concepts like ableism. The COVID-19 pandemic requires bolder action to make sure these voices are included, not silenced.[91] Such bold action necessary to incorporate insights from disabled people's lived experience would require more effective inclusion policies and practices.

First, lived experience has the potential to inform healthcare practice and to help the medical profession to acknowledge, recognise and address the medical ableism that is often presented as scientific objectivity, but risks leaving both patients and practitioners exposed to the harsh consequences of decisions being based on questionable ethical grounds. COVID-19 presents an opportunity to provide all healthcare staff with rapid training on the rights of disabled people.[92] That training should embrace understandings of disability informed by lived experience and should distance itself from categories of 'normal' and 'abnormal' as abstract pathologies.[93] Understanding and awareness of disability ethics can help medical professionals, who have limited knowledge or appreciation of disability experience, when

90 Joseph Stramondo, 'Doing ethics from experience: pragmatic suggestions for a feminist disability advocate's response to prenatal diagnosis' (2011) 4(2) International Journal of Feminist Approaches to Bioethics 48, 71.

91 Emily Lund and Kara Ayers, 'Raising awareness of disabled lives and health care rationing during the COVID-19 pandemic' (2020) 12(S1) American Psychological Association S210, S211.

92 Richard Armitage and Laura Nellums, 'The COVID-19 response must be disability inclusive' (2020) 5(5) Lancet Public Health.

93 Shane Neilson, 'Ableism in the medical profession' (2020) 192(15) Canadian Medical Association Journal E411, E412.

making decisions about how disabled lives are to be valued in a triage situation,[94] on the basis of ethical guidelines.

Second, lived experience can also have a positive effect on health policymaking during the pandemic. For example, one of the policy responses to the pandemic in Northern Ireland has been the proposal to create a Mental Health Champion to represent the views and experiences of patients with a mental health disability and who can hold decision-makers accountable in terms of responding to the ever-changing pandemic landscape. In a statement following the announcement from the Northern Irish Department of Health, Disability Action Northern Ireland (DANI) urged (as a minimum) that criteria for appointment should include that the applicant has personal lived experience of mental ill health.[95]

GDE principles have the capacity to create alternatives to the ethical guidelines. For example, in a recent statement to the BMA following the publication of the BMA ethical guidelines, DANI produced the following guiding principles for ethical guidelines:

> We believe it is critically important healthcare professionals have guidance which includes and accurately reflects disabled people as citizens with fundamental rights (like all others) in the difficult times ahead.
>
> We also believe it is critical that we all have the medical equipment and resources needed.
>
> We call on the BMA to now reach out and meaningfully engage with Disabled People's Organisations. Participation is central to a rights-based approach to health.
>
> We are all in this together.[96]

On the basis of these kinds of principles, guidelines for healthcare decision-making in the context of pandemic-induced scarcity could be altered to express the following. As far as possible, decisions about allocation of scarce healthcare resources should be made in advance and actively include the public (most importantly disabled people themselves).[97] GDE calls for the co-production of guidelines where disability lived experience is an equal in value expertise upon which they

94 Satendra Singh, 'Disability ethics in the coronavirus crisis' (2020) 9(2) Journal of Family Medical Primary Care 167, 171.

95 Disability Action NI, 'Open letter to Minister Swann et al' (DANI, 23 April 2020).

96 Disability Action NI, 'Disability Action deeply concerned by recent BMA Guidance "COVID-19: ethical issues"' (DANI, 2 April 2020).

97 Lawrence Gostin, Eric Friedman and Sarah Wetter, 'Responding to Covid-19: how to navigate a public health emergency legally and ethically' (2020) 50(2) Hastings Center Report 8, 9.

are developed. Guidelines must be transparent and based on clearly explained rationales that are compatible with a person-centred and a human rights-based approach to health.[98] There must be a thorough, individualised review of each patient,[99] grounded in scientific evidence related to transmission of the virus, morbidity and mortality.[100] Such review must avoid explicit or implicit assumptions about the value or quality of life of a patient, based on aspects of their ability unrelated to COVID-19, so that the individual chance of a disabled person with COVID-19 to benefit from treatment is not influenced by how disabled lives are valued by society. Access to treatment decisions should not consider whether someone has a disability, or a proxy for a disability such as 'frailty'. Instead, they should focus on the patient's prospects of benefiting from treatment.[101] Where disabled people have existing health conditions or impairments that are unrelated to their chance of benefiting from treatment, those pre-existing conditions must not play any part in decision-making regarding a disabled person's equal right to access such treatment. ICU triage protocols should focus on identifying the patients who are most likely to die without a ventilator, but are the most likely to survive with one. They should do so using the best available clinical survivability scores, applied on an individual basis, not using broad categorical exclusions.[102] Going further, medical professionals should take decisions based not on an abstract 'norm' of able-bodied (or able of mind), but cognisant that every body (and mind) is different.

Third, and more ambitiously, GDE principles based on lived experience are also a vital part of the developing field of disability ethics. As such, GDE can have the function of transforming practice, through creative and emancipatory disruption of established ways of behaving, established ethical considerations and principles. In the same way in which disability disrupts and challenges ableist normativity, disability ethics can challenge the dubious theoretical grounds, or the uncomfortable compromise between competing ethical frameworks, by ushering in the power of lived experience.

98 Ibid 9.

99 DREDF (n 82).

100 Gostin et al (n 97) 9.

101 Michelle Mello et al, 'Respecting disability rights — toward improved crisis standards of care' (2020) 383 New England Journal of Medicine e26.

102 Mildred Solomon et al, 'Covid-19 crisis triage — optimizing health outcomes and disability rights' (2020) 383 New England Journal of Medicine e27.

CONCLUSION

Grassroots disability ethics principles are those for the ethical and fair distribution of resources and organisation of society that emerge from disabled people themselves. This article has shown that GDE can be used to help illuminate serious problems with COVID-19 guidelines, such as the NICE rapid COVID-19 guidelines on critical care[103] and the BMA 'COVID-19: ethical issues'.[104] Although these guidelines, like many others across the globe, do not overtly discriminate on grounds of disability, they do raise the possibility of indirect discrimination against disabled people, potentially involving the denial of life-saving treatment. More profoundly, the guidelines embody a disableist approach that non-discrimination law alone cannot address. This deeper problem lies with how disabled lives are understood, valued and consequently protected during the pandemic.

More broadly, and perhaps most poignantly, the NICE and BMA ethical guidelines, despite committing to avoiding direct discrimination, may be failing to achieve what they set out to do. The main purpose of both guidelines is to bring clarity and reassurance to both NHS staff who are tasked with making difficult decisions and to disabled patients who have to endure the consequences of these decisions. In their current form, the guidelines provide no such clarity. Instead, they encourage the formation of two opposing and incompatible 'camps': medical professionals versus disabled people, leaving very little space for sharing ideas, experiences and solidarity. Whilst more research is needed into the experiences with regard to the guidelines of both medical professionals and disabled patients during the pandemic, it is nonetheless safe to argue that GDE principles of including disability voices in the legal and policy responses to the COVID-19 pandemic would offer a stronger ethical foundation that brings clarity and reassurance to everyone.

Going further, GDE principles can be used as a foundation upon which to build a disability-inclusive and disability-led response not only to the current COVID-19 crisis, but to other contexts where healthcare resources become scarce. Disability-led narratives on what constitutes ethical, fair and just prioritisation of patients during the pandemic are missing from the guidelines. Yet, these disability-led narratives would offer the key improvement to the guidelines, as they would prevent the historic devaluation of disability to allow for indirect discrimination against disabled patients to be seen as an acceptable means to achieve a legitimate aim. Disability voices and experiences must be included in all policy and legal responses to the current pandemic, as well as

103 NICE (n 2).
104 BMA (n 3).

any future health crisis. Including disability voices and experiences in the construction of legal and policy responses to health crises has the potential to disrupt medicalised 'common sense' on disability in the health field and encourage the cross-pollination of practice with discussions from critical disability studies, disability rights and ethics.

Northern Ireland
Legal Quarterly
Summer Vol. 74 No. 2 (2023) 388–394
Commentaries and Notes DOI: 10.53386/nilq.v74i2.1064

‘Duress of circumstances and voluntary association’: *R v Phair* [2022] NICA 66

John Taggart
Queen’s University Belfast
Correspondence email: john.taggart@qub.ac.uk

ABSTRACT

In the case of *R v Phair*, the Northern Ireland Court of Appeal was tasked with interpreting the scope and application of the comparatively recent criminal defence of ‘duress of circumstances’. While the defence of duress by threats is well established, duress of circumstances has received comparatively little judicial or academic attention. The judgment provides important clarification on the doctrinal and theoretical underpinnings of the defence. Further, the decision is instructive as to how courts should approach the limitation of ‘voluntary association’ which may operate to prevent a defendant successfully pleading the defence.

Keywords: duress; duress of circumstances; duress of threats; necessity.

INTRODUCTION

The defence of duress has existed in English law for centuries.[1] The defence centres on circumstantial pressure and arises where a defendant has completed all of the definitional elements of an offence but, in the circumstances of the case, the defendant’s actions are excused. The defence has most commonly featured in cases where a defendant commits a criminal offence, but does so as a result of threats of death or serious injury. A typical case would involve a defendant being threatened that if they follow orders to carry out some form of assault on another individual, they will be killed. Duress draws heavily on the concept of objectivity, assessing the actions of the defendant against reasonable standards of the ordinary citizen.[2] The defence does not operate to negate a defendant’s *mens rea*, rather it provides an exculpatory excuse to relieve a degree of responsibility for their conduct.

1 Amy Elkington, ‘The historical development of duress and the unfounded result of denying duress as a defence to murder’ (2022) Journal of Criminal Law 1.

2 John Child et al, *Simester and Sullivan’s Criminal Law: Theory and Doctrine* (Hart 2022) 849.

A more recent form of duress to be recognised is that of 'duress of circumstances'.[3] Indeed, the defence appears to have arisen 'more or less by accident' rather than developing separately as a distinct, coherent basis for justifying or excusing criminal conduct.[4] Duress by circumstances essentially covers scenarios where an individual carries out a crime, feeling compelled to do so because of fear arising out of a set of circumstances rather than a person threatening him or her (see further below). Both defences of duress by threats and duress by circumstances are governed largely by the same principles, for example neither is available to a charge of murder or attempted murder.[5]

The defence of duress of circumstances has received relatively little attention from the appellate courts in England & Wales.[6] In an Irish context, the Irish Law Reform Commission (drawing heavily on the experience in England and Wales) has recommended that the defence 'be placed on a statutory footing, having the same scope and application as the defence of duress by threats'.[7] The Northern Ireland Court of Appeal decision in *R v Phair*[8] is valuable in helping us to understand the interaction between the different bases of the duress defence as well as potential limitations placed on it.

BACKGROUND

The appellant was convicted of nine offences after a trial. These included causing death by dangerous driving and causing grievous bodily harm by dangerous driving. The offences related to a fatal car chase which occurred following a failed drugs transaction between the appellant and another man (PT). In short, PT had paid the appellant for cocaine which the appellant did not provide. This resulted in an altercation and a car chase between the two men. The driving of both the appellant and PT was described during the evidence as characterised by speed and

3 The case of *R v Willer* (1986) 83 Cr App R 225 is generally recognised as the first case to demarcate duress by circumstances as a separate offence. This decision was then followed in *R v Martin* [1989] 88 Cr App R 343 in which the court recognised that 'English law does in extreme circumstances recognise a defence of necessity. Most commonly this defence arises as duress that is pressure upon the accused's will from the wrongful threats or violence of another' [345].

4 David Ormerod and Karl Laird, *Smith, Hogan and Ormerod's Criminal Law* (Oxford University Press 2021) 383.

5 Child et al (n 2 above) 849.

6 Karl Laird, 'Duress: *R v Petgrave (Pascoe)* Court of Appeal (Criminal Division): Holroyde LJ, Russell J and HH Judge Mayo QC: 8 June; [2018] EWCA Crim 1397' (2019) 2 Criminal Law Review 160.

7 Law Reform Commission (Ireland), 'Report: Defences in criminal law' (LRC 95-2009, December 2009).

8 [2022] NICA 66.

'a hot pursuit'.[9] The appellant, who was driving one of the cars, was injured. His girlfriend was killed, and another young woman was also seriously injured. A set of facts was agreed between the defence and prosecution. It was agreed that the death of the appellant's girlfriend was caused by the injuries she sustained in the vehicle collision. It was also agreed that at the time of the accident the appellant's blood contained Alprazolam, also known by the brand name Xanax, well above the range expected following therapeutic use.

The appeal was based on five principal grounds which included the decision to place evidence of the appellant's bad character before the jury, admission of hearsay evidence against the appellant and the placing of a limitation on the defence of duress of circumstances. In relation to the third ground, there was no dispute that the defence of duress was properly left to the jury. However, it was contended that the judge should not as a matter of law have included a voluntary association limitation as part of his direction. In the alternative the appellant argued that there was no evidential basis for the limitation of the defence to be left to the jury.

Under the sub-heading 'defence of duress', the trial judge set out for the jury a route to verdict. It was the final element of this route, which outlined the 'voluntary association limitation', that formed the basis of the appellant's case. The core question was whether the inclusion of this limitation was correct in law where there were no direct threats which compelled the appellant to commit crimes but, rather, he committed crimes due to the circumstances that arose. This third question was framed by the trial judge as follows:

> Had the defendant voluntarily put himself in a position in which he knew or ought reasonably to have known that he might be compelled to commit a crime by threats of violence made by other people?
>
> If you are sure that this was the case the defence of duress is not available and you should return a verdict of guilty.
>
> If you are sure that this was not, or you think it may not, have been the case, you should find him not guilty.

The Court of Appeal declined to analyse the nature of duress or engage in any academic debate as to the relationship between this defence and necessity. Instead, the court narrowed its focus on the question of 'whether the defence having arisen, the jury should also have been told that it was not available if the appellant had voluntarily exposed himself to the risk of compulsion to commit crimes'.[10] The court surveyed the

9 Ibid [6].
10 Ibid [62].

relevant caselaw, starting with the seminal authority of *R v Hasan*.[11] It emphasised the point made by Lord Bingham that a defendant may not rely on duress to which he has 'voluntarily laid himself open'.[12] The rationale for this limitation was based on the imperative of discouraging association with known criminals, and that the law should be 'slow to excuse the criminal conduct of those who do so'.[13] Further echoing the judgment of Lord Bingham in *Hasan*, the Northern Ireland Court of Appeal recognised that the net of voluntary association is cast wide and that it is not confined to foresight of coercion to commit crimes.[14] The Court of Appeal further cited the case of *R v Ali*[15] as authority for the notion that it is 'the risk of being subjected to compulsion by threats of violence that must be foreseen or foreseeable that is relevant, rather than the nature of the activity in which the threatener is engaged'.[16]

The appellant's legal representatives submitted that there are logical public policy and moral distinctions between the two different forms of duress. It was argued that the voluntariness limitation should not apply in circumstances where an individual commits a criminal offence in escaping a threat of death or serious injury from an associate. Public policy, it was argued, should not be so broad as to mean that a criminal associate can never rely on duress of circumstances where they are in the act of attempting to escape from a threat of death or serious injury. The Northern Ireland Court of Appeal was unconvinced by these arguments and refused to draw a distinction between the two forms of the defence. The court sought to highlight commonalities between the two forms of duress and noted that:

> ... in drawing all of the above strands together it is our view that the voluntary association limitation is not confined to circumstances of direct threat. It seems to us proper to apply it in other circumstances where the threat is implied or derived from circumstances i.e., duress of circumstances. It would be artificial and against public policy to make a distinction.[17]

The court therefore dismissed this ground of appeal, saying that in the circumstances the appellant could have foreseen or ought reasonably to have foreseen the risk of being subjected to compulsion to act in a criminal way by threats of violence to commit criminal offences. In the words of Lord Bingham in the case of *Hasan*, the appellant was unable to rely on the defence of duress to excuse '*any act*' (emphasis added)

11 [2005] 2 AC 467.
12 Ibid [21].
13 Ibid [38].
14 *Phair* (n 8 above) [72].
15 [2008] EWCA Crim 716.
16 *Phair* (n 8 above) [74].
17 Ibid [84].

which he was compelled to do.[18] In the present case, that included causing death or grievous injury by dangerous driving. Each of the other grounds of appeal were also dismissed.

COMMENTARY

As a general rule, an individual who voluntarily accepts the risk of being placed in the 'do it or else' dilemma is not permitted to use that dilemma as an excuse within the law of duress (although it may amount to mitigation in some circumstances). There are strong policy reasons behind such an approach: the law should discourage association with known criminals and any consequences which arise from such association will very rarely be excused.[19] As Lord Lowry LCJ remarked in the seminal authority of *R v Fitzpatrick*, a defendant could not be allowed to 'put on, when it suits him, the "breast plate of righteousness"' by raising duress in an effort to escape criminal liability.[20] The appeal in *Phair* concerned the breadth of that rule and how it applies to the different forms of duress (duress of threats and duress of circumstances). The facts of the case did not involve the typical duress by threats scenario, rather the appellant argued that he was impelled to act as he did because, based on what he reasonably believed the situation to be (ie being pursued by a car at speed), he had good reason to fear death or grievous bodily harm would result.

Several interesting points arise from the Court of Appeal's analysis. Firstly, the court did not enquire into the precise relationship between duress of circumstances and necessity as a defence. Indeed, in the case of *R v Conway* the England & Wales Court of Appeal collated the two and concluded that no relevant doctrinal differences exist.[21] In other scenarios, however, the difference between the two may attain more significance. The defence of duress focuses on whether a person of reasonable firmness would have acted similarly, while necessity asks whether, in the circumstances the defendant was in (or reasonably believed themselves to be in), it was legitimate to break the law. As Simester and Sullivan point out, the latter question does not always require an emergency or involve an imminent threat.[22] Furthermore, necessity has traditionally been regarded as a justificatory defence while duress of circumstances has been construed as excusatory

18 *Hasan* (n 11 above) [38] (Lord Bingham).
19 Ibid.
20 [1977] NI 20 [31].
21 [1989] QB 290 [297].
22 Child et al (n 2) 866.

in nature by courts in England and Wales.[23] Interestingly, recent academic commentary has analysed doctrinal similarities between the defences in the wake of the Covid-19 pandemic.[24]

The arguments advanced by the appellant on this appeal point are worth considering. Essentially, the appellant's representatives sought to distinguish the type of case where the defendant 'complies' with criminal associates under threat and cases where the same person commits a criminal offence in 'escaping' a threat of death or serious injury from an associate. It is submitted that the Northern Ireland Court of Appeal was correct to reject this conceptualisation and defer to the wider concept of voluntary association as outlined by Lord Bingham in *Hasan*. The case law in this area is underpinned by the notion that association with known criminals is sufficient to disqualify a defendant from relying on the defence.[25] As such, the 'escape' distinction advanced by the appellant in this case could not find support due to the prior association with known criminals. In theory, no attempted escape from associated known criminals could ever permit an 'escape' argument to succeed. It is interesting that later in the judgment the Court of Appeal was content to frame the appeal as an 'escape case' as opposed to one of self-defence.[26] As such, convincing the court that a defendant was escaping a threat of death or serious injury from an associate will itself never be sufficient to ground a defence of duress by circumstance. Indeed, in both cases involving defences of duress by threats and duress of circumstances, a relevant consideration for the jury will also be whether the defendant took reasonable steps to escape from the threat faced. The line between these two forms of escape will likely be highly fact-dependent.

The outcome of the appeal in *Phair* serves to highlight just how important voluntary association will be on any attempt to run a duress defence. As outlined by the England & Wales Court of Appeal in *R v Harmer*, the prosecution must demonstrate no more than the fact that the defendant voluntarily exposed himself to unlawful violence.[27] There was, the England & Wales Court of Appeal held, no further requirement that the defendant foresaw that he might be required under the threat of violence to commit crimes.[28] Interestingly, the third and

23 William Wilson, *Criminal Law* (Pearson Education 2014) 263; Christopher Clarkson, 'Necessary action: a new defence' (2004) Criminal Law Review 81.

24 See, for example: Bill Clawges, 'Reexamining the application of duress and necessity defenses to prison escape in the context of COVID19' (2022) 112 Journal of Criminal Law and Criminology Online 83.

25 Child et al (n 2 above).

26 *Phair* (n 8 above) [131].

27 [2002] Crim LR 401.

28 Ibid [16]–[17].

final question in the trial judge's 'route to verdict' in the present appeal required the jury to ask themselves if the appellant 'ought reasonably to have known that he might be compelled to commit a crime by threats of violence made by other people'.[29] As has been explained above, this in fact puts the prosecution's task too high. Association with known criminals and the related exposure to unlawful violence is sufficient for disqualification of the defence.

Finally, the Court of Appeal was tasked with settling a disparity between law and practice in England & Wales and Northern Ireland. It noted that some 'confusion has arisen by virtue of the Crown Court Bench Book NI which does not specifically provide for a voluntary association limitation being applied to a defence of duress by circumstance'.[30] This position, the Court of Appeal pointed out, was at odds with the Crown Court Compendium in England & Wales and indeed the established caselaw. For example, Lord Woolf LJ in *R v Conway* when examining the parameters of duress of circumstances remarked that 'what is important is that, whatever it is called, it is subject to the same limitations as the do this or else species of duress'.[31] While this omission in the Crown Court Bench Book NI was likely a mere oversight, the Court of Appeal's realignment of the Northern Irish law with England & Wales emphasises the need for coherence and consistency in this area. While duress of threats and duress of circumstances are recognised as separate defences, there is clearly an imperative to recognise their common elements. As Ormerod and Laird have noted, duress of circumstances has developed by analogy to duress by threats, and, as such, there is a 'ready made set of principles to govern it'.[32] The Northern Ireland Court of Appeal decision in *Phair* pays regard to this reality and aids understanding of the operation of duress of circumstances in practice.[33]

29 *Phair* (n 8 above) [58].
30 Ibid [76].
31 [1989] QB 290 [297].
32 Ormerod and Laird (n 4 above) 363.
33 Clarkson has proposed a different direction for the defences of duress, namely advocating for a collapsing of the defences of duress by threats and circumstances, necessity and self-defence into one general defence of necessity, termed 'necessary action'. See Clarkson (n 23 above).

Northern Ireland
Legal Quarterly
Summer Vol. 74.2 (2023) 395–415
Commentaries and Notes DOI: 10.53386/nilq.v74i2.1100

From the Protocol to the Windsor Framework

C R G Murray
University of Newcastle
Niall Robb*
Queen's University Belfast
Correspondence email: colin.murray@newcastle.ac.uk; nrobb03@qub.ac.uk

ABSTRACT

The Windsor Framework, the new package of measures agreed by the United Kingdom (UK) and European Union (EU) as well as the new name for the Protocol on Ireland/Northern Ireland, was presented in February 2023 amidst considerable fanfare. This article examines the rationale for the new Framework amongst the negotiators and how some of its headline provisions impact upon those most exposed to the out-workings of any deal – those living and doing business in Northern Ireland. We investigate the possible implications for Northern Ireland of the new minimalist regulatory alignment in the trade in goods and the possibility of a 'cooperation dividend' stemming from warmer UK–EU relations. In particular, we examine the operation and possible limitations upon the 'Stormont Brake' mechanism. This article ultimately assesses whether Sunak's Windsor Framework will be any more successful than the May Backstop and Johnson Protocol before it at 'getting Brexit done'.

Keywords: Windsor Framework; Brexit; Northern Ireland; trade; regulation.

TOWARDS FIXING THE PROTOCOL

It takes a long time for the fury and animosity generated by an upheaval like Brexit to subside, especially after seven years of repeatedly traversing the infeasibility of imposing a customs and regulatory border across the island of Ireland and attempting multiple variations of an alternative to doing so. The Windsor Framework, met with widespread acclaim at Westminster and evident bonhomie between Rishi Sunak and Ursula von der Leyen, could nonetheless

* Respectively, Professor of Law and Democracy, Newcastle University, and PhD Candidate, Queen's University Belfast. Our thanks to Steve Peers (University of Essex) for his advice on this article, which has been adapted from C Murray, 'The Protocol deal: a "just like that" moment seven years in the making' (*DCU Brexit Blog* 28 February 2023). Article updated and hyperlinks last accessed on 17 April 2023. This article builds upon the content of the 2022 special issue of the *Northern Ireland Legal Quarterly*, 'Northern Ireland's Legal Order after Brexit'.

mark the point at which we move into a genuinely post-Brexit phase of relations between the European Union (EU) and the United Kingdom (UK). This leaves the question, beyond the prominent rebranding, of whether Sunak's Windsor Framework will prove any more successful than the May Backstop and the Johnson Protocol before it.

In headline terms, under the Windsor Framework Northern Ireland remains aligned with the EU Single Market rules for goods, and as a result the role of the Court of Justice of the European Union (CJEU) in the oversight of those rules also remains. Although the Windsor Framework may not present a new model for the goods arrangements covering Northern Ireland, it nonetheless provides for a significant reworking of the Protocol, primarily effected through Joint Committee action and mutually agreed unilateral measures by the UK and the EU. For Sunak, this overhaul of the Protocol is so comprehensive that it has 'removed the border in the Irish sea'.[1] In assessing the Windsor Framework, this article explores why the interested parties might regard these prominent elements of the new deal as attractive and proceeds to explore some of the most significant claims made about the new deal in depth. These include the changes to the goods rules applicable to Northern Ireland, the position of Northern Ireland with regard to new trade agreements concluded by the UK, the new Northern Ireland Assembly mechanism for objecting to dynamic alignment with aspects of EU law (the 'Stormont Brake'), and the deal's impact on Northern Ireland's participation in the Withdrawal Agreement's committee structures.

These changes are brought about by a complex combination of legal mechanisms using variation provisions contained in the Withdrawal Agreement itself. Using article 164 as their legal basis, the UK and EU have made full use of the considerable leeway this provision provides them to 'address omissions or other deficiencies, or to address situations unforeseen when this Agreement was signed'[2] during the first four years following the conclusion of the Withdrawal Agreement's transition period. These terms do not, however, allow the Joint Committee to 'amend the essential elements of the Agreement', and the Commission has found itself at the edges of its negotiating mandate in concluding the Windsor Framework. The deal has therefore required a combination of unilateral commitments by the UK and the EU (the latter providing carve-outs for Northern Ireland from aspects of EU law hitherto operative under the Protocol, modelled upon the EU's

1 R Sunak, HC Deb 27 February 2023, vol 728, col 572.

2 Agreement on the Withdrawal of the United Kingdom of Great Britain and Northern Ireland from the European Union (30 January 2020), art 164(5)(d).

unilateral reforms to medicines regulations implemented in 2022)[3] and Joint Committee decisions and declarations.

DOES BETTER MEAN BETTER FOR EVERYONE?

The Windsor Framework provides a policy victory for many of the key stakeholders involved in determining Northern Ireland's post-Brexit arrangements. From the perspective of the UK Government, Rishi Sunak needs a major achievement to convince voters of his effectiveness as leader if he intends to claw back the Conservatives' position in the polls following Liz Truss's short and turbulent premiership. Having evidently concluded that he could not risk a trade war with the EU amid a cost-of-living crisis, he was willing to gamble that, for all the manoeuvrings ahead of the deal announcement, all but the most stalwart European Research Group (ERG) members within his parliamentary party would balk at the prospect of bringing down his Government at a time when the polls suggested that a general election would be disastrous for the Conservative Party. Sunak's Government can credibly argue that its more constructive approach to UK–EU relations, after the turbulence of the Johnson era, has led the EU to make concessions over several controversial aspects of the Protocol.

For the Democratic Unionist Party (DUP), whose opposition to the Protocol provided the pretext for the UK Government's refusal to fully implement its terms, the Prime Minister will hope that it will ultimately accept the EU's tangible movement on issues including pet movements, medicines, regulation, customs and value-added tax (VAT) (all of which directly address consistent DUP talking points) and restore power-sharing. It is unlikely that holding out will yield further concessions, and prolonging Northern Ireland's instability is not going to strengthen its place in the Union in the long term. Even the deal's title has been calibrated to apply the soothing balm of monarchism to the DUP's wounded sensibilities.[4] Moreover, now that the UK Government has made arrangements to manage the governance of Northern Ireland pretty much indefinitely by a form of quasi direct rule, the DUP's future leverage is uncertain.[5] Rather, by returning to

3 Directive 2022/642/EU amending Directives 2001/20/EC and 2001/83/EC as regards derogations from certain obligations concerning certain medicinal products for human use made available in the United Kingdom in respect of Northern Ireland and in Cyprus, Ireland and Malta.

4 The deal was concluded at a hotel near Windsor Great Park; E Ng, 'Fairmont Windsor Park Hotel delighted to be part of "historic occasion"' *The Independent* (27 February 2023).

5 Northern Ireland (Executive Formation etc) Act 2022 and Northern Ireland (Executive Formation and Organ and Tissue Donation) Act 2023.

power-sharing, the DUP can employ the Windsor Framework's new mechanisms to resolve issues – the Framework being a mechanism for managing the UK–EU relationship with regard to Northern Ireland as much as a solution for particular issues. Only through taking its place within the Northern Ireland Executive will the DUP secure its say in the Withdrawal Agreement's committee structures and only through a functioning Assembly can the new 'Stormont Brake' operate.

The Windsor Framework, however, remains a reformulation of the Protocol rather than a new model, which does not address the party's persistent complaint that article VI of the Act of Union had been undermined.[6] This 'constitutional' objection might possess much less traction at Westminster since the UK Supreme Court comprehensively dismissed the *Allister* challenge to the constitutionality of the existing Protocol arrangements,[7] but it has become very difficult for the DUP's leadership to disavow.[8] Any effort to set aside these concerns and back the deal risks exacerbating the party's internal divisions over the shape of Brexit and could enable some of the voices that the party has encouraged within Loyalism, in order to reinforce opposition to the Protocol, to achieve even greater prominence within Unionist discourse. A further challenge for Sir Jeffrey Donaldson, if he were to attempt to change tack, is finding a window between successive looming elections to make any *volte face*. These circumstances saw the DUP vote against the statutory instrument implementing the UK's Windsor Framework commitments when it came before Parliament in March 2023, but, in the absence of a major Eurosceptic push against Sunak, this did not disrupt the measure becoming law (discussed below). The door remains open for the DUP to agree, however grudgingly, to participate under the Framework, and thereby attempt to use the deal's new mechanisms to challenge aspects of Northern Ireland's alignment with EU law.

On the other side of the table, the European Commission might well simply be questioning the point of continuing interminable struggles over the precise rules applicable to Northern Ireland, unimportant as it is in the context of the size of the EU Single Market. Michel Barnier's memoirs attest to the grip of Brexit 'fatigue' as early as 2019.[9] Having secured improved data-sharing from the UK Government, allowing it to conduct a managed-risk approach to the Protocol, the Commission may feel secure enough to accept mitigations to the Protocol's terms. Ireland, however, seems to have conceded most through this deal

6 J Donaldson, HC Deb 27 February 2023, vol 728, col 581.

7 *In re Allister* [2023] UKSC 5, [66]–[68] (Lord Stephens).

8 J Webber, 'Northern Ireland unionist leader demands "change" to Windsor Framework' *Financial Times* (London 14 March 2023).

9 M Barnier, *My Secret Brexit Diary: A Glorious Illusion* (Polity Press 2022) 217–222.

(even if it was not directly at the table). The trade dislocations inherent in the Protocol led traders to explore North–South opportunities due to Brexit barriers to trade.[10] Mitigate those barriers and much of this new business will likely fail to materialise. Perhaps the calculation is that the deal does enough to protect existing trade and that damage to North–South co-operation and UK–Ireland relations at a political level meant that attempting to hold onto these potential benefits was no longer worth the effort. The requirement of distinct labelling for some products to be sold in Northern Ireland, as 'Not for EU', could nonetheless generate ongoing difficulties for businesses which regard a small market as not being worth differentiated packaging,[11] and it will be difficult to assess the impact of this factor upon supply chains until the system is operative.

The conclusion of the Windsor Framework is thus a function of the most significant actors having finally exhausted their energy around this most intractable aspect of Brexit and seeking to reset relations. This is not to say that the substance of the deal is unimportant; Northern Ireland is a small, peripheral and underdeveloped economy.[12] Anything in the Windsor Framework that helps to alleviate barriers to trade and stimulate economic regeneration is therefore particularly important in light of these pressures, but there was no obvious way to reach these new arrangements without having gone through the fraught negotiations to date. A Protocol which established the basic parameters for trade in goods and regulation with regard to Northern Ireland had to be introduced and tested before it could be mitigated, even if that left people and businesses in Northern Ireland feeling like the subject of an experiment. All the tortured steps and missteps of recent years have involved the UK and EU working out what they can accept as a tolerable post-Brexit relationship; when Northern Ireland required prominent and complex accommodations, repeated recalibration of these arrangements was an inevitability.

---

10 S Lowe, 'Six sides to every story: trade in goods and the Northern Ireland Protocol' (2022) 17 Journal of Cross Border Studies in Ireland 77, 87.

11 The UK Government has rebuffed concerns with statements that these new requirements amount to '[p]roportionate' and 'phased' arrangements: L Docherty, HC Deb 20 March 2023, WA 164046.

12 See G Brownlow, 'Northern Ireland and the Economic Consequences of Brexit: taking back control or perpetuating underperformance?' (2023) Contemporary Social Science (Advanced Access).

A MINIMALIST TAKE ON NORTHERN IRELAND–EU REGULATORY ALIGNMENT

For Sunak, the most pressing issue with the Protocol has been how it 'treated goods moving from Great Britain to Northern Ireland as if they were crossing an international customs border'.[13] The Windsor Framework offers an alternative to the Protocol's previous model of alignment in light of the concerns raised by politicians, business and civil society in Northern Ireland. Most notably, the Windsor Framework permits the partial disapplication of EU rules for goods, provided their final destination is in Northern Ireland. This is highlighted in the Joint Committee decision amending article 6(2) of the Protocol. It includes the following text, which recognises specific implications of Northern Ireland's place in the UK's internal market:

> This includes specific arrangements for the movement of goods within the United Kingdom's internal market, consistent with Northern Ireland's position as part of the customs territory of the United Kingdom in accordance with this Protocol, where the goods are destined for final consumption or final use in Northern Ireland and where the necessary safeguards are in place to protect the integrity of the Union's internal market and customs union.

This acknowledgment enables a reworking and streamlining of the applicable EU rules under the Protocol and facilitates the creation of a 'green lane' through which goods travelling to Northern Ireland are moved with greater ease than those which are at risk of onward movement into the EU.[14] The latter will transit the Irish Sea on the basis of 'red lane' arrangements, to be applied at the Irish Sea border as if the goods were entering the Single Market from an external country.

The development of a differentiated regime for goods moving between Great Britain and Northern Ireland, dependent on their end destination, has been the most obvious overlap between the UK and EU plans for Protocol reform since the Commission published its proposals in October 2021.[15] The difficulty has been defining what goods are covered by the green lane arrangements and establishing the necessary safeguards. The new arrangements, in the form of a new Joint Committee determination of goods at risk of onward movement, are focused in large part on goods subject to sanitary and phytosanitary (SPS) checks, alongside products to be used in construction in Northern Ireland.[16] In doing so, the EU has agreed

13 Sunak (n 1 above) col 570.

14 Joint Committee Decision 1/2023, art 7.

15 C Murray, 'From oven-ready to indigestible: the Protocol on Ireland/Northern Ireland' (2022) 73(S2) Northern Ireland Legal Quarterly 8, 31.

16 Joint Committee Decision 1/2023, arts 6 and 7.

to a risk-based approach to checks on goods and the UK has accepted the data-sharing procedures, labelling requirements and the border infrastructure necessary to facilitate this.[17] The new arrangements provide what Sunak describes as:

> [A] new, permanent, legally binding approach to food. We will expand the green lane to food retailers, and not just supermarkets but wholesalers and hospitality, too. Instead of hundreds of certificates, lorries will make one simple, digital declaration to confirm that goods will remain in Northern Ireland.[18]

SPS checks are particularly extensive under EU law and have generated some of the most persistent concerns over the eventual implementation of the Protocol, leading the UK Government to unilaterally extend the relevant grace periods throughout its operation to date.[19] With issues like gene editing being a prominent part of the UK Government's policy agenda, the agrifood context is likely to be subject to significant standards divergence.[20] The Windsor Framework does not do away with these challenges, instead it substitutes packaging requirements and trade-flow oversight for systematic compliance checks. This arguably substitutes one non-tariff barrier to trade for another, with the new arrangements being less onerous for hauliers but imposing more on businesses in Great Britain making food products for retailers in the small Northern Ireland market.

The tailoring of the green lane to address this specific issue, moreover, narrows its operation. These new arrangements do not, for example, apply to manufactured goods. But accompanying the specific rules applicable to SPS checks, there is a general green-lane exception for Northern Ireland businesses with turnovers of under £2 million bringing goods from Great Britain into Northern Ireland, which the UK Government has presented as 'meaning four-fifths of manufacturing and processing companies in Northern Ireland who trade with Great Britain will automatically be in scope'.[21] These new arrangements therefore address some of the established frictions at the sea border, but do so in a way that channels trade through business entities, in an effort to constrain abuses. As some have noted, this business-centred

17 Ibid arts 9–12

18 Sunak (n 1 above) col 571.

19 C Barnard, 'The status of the Withdrawal Agreement in UK Law' in C McCrudden (ed), *The Law and Practice of the Ireland–Northern Ireland Protocol* (Cambridge University Press 2022) 107, 113.

20 Protocol on Ireland/Northern Ireland Sub-Committee, 'Oral Evidence: The Windsor Framework' (22 March 2023) Q5 (Jess Sargeant).

21 HM Government, *The Windsor Framework: A New Way Forward* (2023) CP 806, para 12.

approach does little for consumers seeking to buy products direct from businesses in Great Britain.[22]

The Windsor Framework does not therefore so much 'remove' the Irish Sea border, as Sunak has claimed, as address a series of problems with its operation which had already been identified and which have generated concerns within Northern Ireland. Its green-lane provisions do not, of themselves, prevent those problems re-emerging, particularly for larger companies, should the UK Government seek to put in place product standards which diverge from those applicable under the EU Single Market.[23] The expansion of the scope of the general 'by turnover' exemption will nonetheless mean that any such burdens will not fall heaviest on businesses least able to deal with the associated costs. These risk-based arrangements, moreover, place a new focus on the land border between Ireland and Northern Ireland in circumstances where risks are identified. As the UK Government's Command Paper explains, 'checks North–South on the island of Ireland will … operate on a risk and intelligence-led basis'.[24]

This new approach has also seen UK Government ministers insist that the deal enables 'free-flowing trade within the whole United Kingdom', including Northern Ireland's full participation in new UK trade agreements.[25] Under the Protocol, for all that Northern Ireland was stated to be part of the UK's customs territory, the arrangement carried with it the obligation to apply EU customs and regulatory rules for goods, and in some cases EU tariffs, at ports of entry into Northern Ireland. Under the Windsor Framework, goods can be moved into Northern Ireland (as an end destination) using the green lane process under the conditions discussed above, but with the additional proviso, relevant to UK trade agreements, that 'the duty payable according to the Union Common Customs Tariff is equal to zero'.[26] Given the scope of the Trade and Co-operation Agreement, tariffs are zero for goods moving between Great Britain and the EU where either is the point of origin of the goods. This is not, however, the case where the UK enters a trade agreement with a third country with more generous terms for certain goods than the EU. There are, however, exceptions to this requirement for trusted traders moving goods for final use in Northern Ireland and parcels. The extent to which these exceptions will be effective in allowing Northern Ireland businesses with turnovers of

22 See S McBride, 'Buying GB plants online seems set to be banned under EU deal … and NI Secretary misled public' *Belfast Telegraph* (18 March 2023).

23 A McCormick, 'The Windsor Framework: a quick evaluation' (*UK in a Changing Europe* 28 February 2023).

24 HM Government (n 21 above) para 46.

25 Sunak (n 1 above) col 570.

26 Draft Joint Committee Decision 1/2023, art 7(1)(a)(i).

under £2 million to participate in new UK trade deals thus depends on whether the Commission is concerned about leakage of goods into the Single Market using the green lane and seeks that checks be imposed as a result.

Other significant changes to the Protocol relate to VAT and excise and a range of unilateral EU measures on parcels, agri-food, plants and pets, and medicines. Under the Windsor Framework, VAT rates on some products remaining in Northern Ireland can now be varied and a process is foreseen which will expand the areas to which this applies. The special arrangements for agri-food, plants, pets, and medicines are implemented through carve-outs to EU legislation in terms of goods movements to Northern Ireland or sale there, which further cements Northern Ireland's unique status with regard to the EU legal order. Indeed, the UK Government went so far as to claim that 1700 pages of EU law will no longer apply to Northern Ireland as a result of the deal, but this has been a headline that ministers have struggled to elaborate upon, and which only applies to the specific application of these rules to certain trading arrangements.[27] These elements of the deal have been designed to cut the amount of EU law required for goods regulation in Northern Ireland to the minimum required 'to maintain the unique ability for Northern Ireland firms to sell their goods into the EU market'.[28] This enables the UK Government to trumpet a reduced role for the CJEU in Northern Ireland as part of the Windsor Framework,[29] even though the nature of its role remains unchanged because there will be less applicable EU law to require its input.

The impetus behind the deal is addressing specific, identified problems with the Protocol, rather than attempting to reconsider its overarching operation. This is not unimportant; the EU has long been presented by the UK Government as rigid and doctrinaire in its application of Single Market rules. It therefore appears to be prioritising 'pragmatic' solutions to the problems generated by the Protocol.[30] This makes it impossible to quantify, as a bald figure, the amount of EU law that no longer applies in Northern Ireland as a result of the Windsor Framework; different amounts apply depending on whether the green lane, the red lane or rules applicable to production in Northern Ireland are invoked, as one minister has been obliged to acknowledge:

---

27 HM Government (n 21 above) para 8. The Secretary of State for Northern Ireland has struggled to account for this and related claims: European Scrutiny Committee, 'Oral Evidence: Government Northern Ireland Protocol negotiations' (2023) HC 1101, Q104–109, Q119 (Chris Heaton-Harris).

28 Ibid para 57.

29 Ibid para 29.

30 S Peers, 'The Windsor Framework: limiting the scope of EU law in Northern Ireland in practice, though not in theory' (*EU Law Analysis* 4 March 2023).

> I am not an expert in EU law, and I have no intention of becoming one, but my understanding is that the situation is somewhat more complex than just adding together a list. There will of course be some directives that are in part still applied, in respect, for example, of the red channel, and disapplied in respect of the green channel.[31]

The problem, however, is that Single Market rules are much less coherent when hedged with exceptions. In its eagerness to highlight how the deal addresses problems affecting trade across the Irish Sea, the UK Government's account of the Windsor Framework presents Northern Ireland as akin to an outsized Freeport, and not as a space in which goods are produced and consumed.[32] As a result, Northern Ireland is left with a somewhat moth-eaten set of EU rules applicable to trade in goods from Great Britain (provided the goods' ultimate destination is Northern Ireland) and a rather more extensive list of EU rules applicable to the production or processing of goods in Northern Ireland. This could see goods in Northern Ireland being produced to higher standards than those required in Great Britain, and facing competitive disadvantages within Great Britain as a market as a result.[33] The Assembly and Executive, if reconstituted following this deal, will be presented with the challenge of making these complex sets of rules work amid any push for divergence of rules by the current UK Government. All of which will put considerable pressure on how Northern Ireland's institutions approach alignment with EU law.

THE 'STORMONT BRAKE'

If the red and green lane system was an expected element of the 'landing zone' for a deal, and the UK Government has accepted a continued role for the CJEU, the Windsor Framework's attempt to address the Protocol's democratic deficit[34] – the 'Stormont Brake' – is the major flourish within Sunak's offering. He presented this to Parliament as a means by which the Northern Ireland Assembly can 'block' dynamic alignment, and that, if this mechanism were to be employed, 'the UK Government will have a veto' ending 'the automatic ratchet of EU law'.[35] These claims require some unpacking, not least because of the slippage in Sunak's pledges between an Assembly 'block' and a UK

31 Lord Caine, HL Deb 7 March 2023, vol 828, col 686.

32 V Gravey and L C Whitten, 'The Windsor Framework and the environment' (*Brexit and Environment* 7 March 2023).

33 See A Bounds and J Webber, 'New EU arsenic rules catch Northern Ireland between Brussels and London' *Financial Times* (London 9 March 2023).

34 See A Deb, 'Parliamentary sovereignty and the protocol pincer' (2023) 43 Legal Studies 47, 63–64.

35 Sunak (n 1 above) col 574.

Government 'veto'. There is further complication in the Command Paper, accompanying the Windsor Framework, presenting these measures as requiring cross-community consent.[36]

The arrangements, incorporated into a new article 13(3a) of the Protocol as amended by the Windsor Framework, with further details set out in a Unilateral Declaration by the UK (and in subsequent enacting legislation, discussed below),[37] are subject to several conditions. The first relates to the new EU law in question. The Brake does not apply to changes in all EU law applicable in Northern Ireland under the Protocol, but only to measures covered in parts of annex 2 of the Protocol (relating to the single market in goods). This means that automatic dynamic alignment will still, for example, apply to the EU Equality Directives listed in annex 1, which underpin much of the operation of article 2 of the Protocol.[38] Where an applicable EU law development is at issue, its content or scope must 'significantly' differ from the measure it is replacing or amending, generating 'significant impact' for everyday life in Northern Ireland.

The second condition relates to Northern Ireland's power-sharing institutions. The Executive and Assembly must be functioning; no Stormont, no Stormont Brake. This is not just a carrot to tempt the DUP back into power-sharing, rather it is a further precondition of the Brake being used that the Assembly and the Executive have consulted on the measures with the public in Northern Ireland and also availed of the consultative mechanisms in place with the EU and the UK Government. If this has been done, the operative part of the Brake draws upon part of the Petition of Concern mechanism, as modified following the *New Decade, New Approach* deal to restore power-sharing in 2020.[39] It allows a group of 30 of Stormont's 90 Members of the Legislative Assembly (MLAs) from two parties to notify the UK Government of their desire that the Brake be applied. The Stormont Brake is not necessarily, therefore, a DUP veto; no party has won 30 or more seats since the Assembly was reduced in size from 108 to 90 MLAs.[40] To meet the requirements for the Brake therefore, the DUP, currently with 25 MLAs, would have to cooperate with the Ulster Unionist Party. Other partnerships would either fall short of 30 MLAs or would be unlikely to share the DUP's concerns with regard to particular EU laws.

---

36 HM Government (n 21) para 64.

37 Draft Joint Committee Decision 1/2023, annex I Unilateral Declaration by the United Kingdom: Involvement of the Institutions of the 1998 Agreement.

38 Draft Joint Committee Decision 1/2023, art 2.

39 *New Decade, New Approach* (8 January 2020), para 12.

40 Northern Ireland (Miscellaneous Provisions) Act 2014, s 6.

Under the terms of article 13(3a), however, this is not a cross-community measure, but rather a simple counter-majoritarian measure; indeed, given Sinn Féin's historic reticence towards EU law, there might in future be common (majoritarian) cause over the impact of particular EU law developments upon Northern Ireland. The UK Government has, however, suggested additional cross-community requirements in its enactment of domestic measures to give effect to the Brake, reflecting some of the language found in the Command Paper but not in the Windsor Framework documents relating to the Brake.[41] This suggests that there may have been different understandings of how the Brake works and its interaction with the Petition of Concern mechanism within the UK Government, and the implementation legislation (discussed below) is being viewed as a means of smoothing out these discrepancies. Any such finessing, however, is difficult to square with the terms of the UK's Unilateral Declaration, which opens with a commitment to 'adopt the following procedure to operate the emergency brake mechanism in Article 13(3a) of the Windsor Framework'.[42]

A third condition relates to the timing of the Brake's employment. All of these processes must take place 'within two months of the publication of the specific Union act' within the Official Journal of the European Union. This will normally be early in the measure's transposition period (where relevant) and poses a challenge given the consultation requirements which must be followed to use the Stormont Brake. The Assembly parties will thus have to carefully assess EU measures during the EU legislative process and be ready in advance of a new measure's publication to engage the Brake. This will, of course, be easier if they are fully engaged in the Withdrawal Agreement's consultative mechanisms, and the expectation is likely that this engagement will head off disputes, potentially with Northern Ireland-specific adjustments being made to EU law as it is drafted. Careful collaboration between the Northern Ireland Executive, the European Commission and the UK Government will be required to enable Northern Ireland's institutions to stay abreast of developments and consider emerging EU law under the scope of the Brake. The good functioning of the new Windsor Framework Democratic Scrutiny Committee to be formed in the Assembly pursuant to the implementing Statutory Instrument (discussed below) will be vital to monitoring and scrutinising this legislation.[43] Otherwise, the risk is that the full

41 Lord Caine, HL Deb 2 March 2023, vol 828, col 381.

42 Draft Joint Committee Decision 1/2023, annex I (n 37 above) para 1.

43 Windsor Framework (Democratic Scrutiny) Regulations 2023 (SI 2023/XX) para 2.

implications of an EU measure might not be recognised until too late into the transposition process to prevent it coming into force.

Having triggered the Brake, the Stormont parties taking issue with a new EU law will have to convince the UK Government that they have met all conditions and of their substantive case as to the problems generated by the EU measure in question. Even then, the Brake can only be invoked with regard to the specific parts of the EU law over which the problems are demonstrated; if only parts of the measure cause significant identifiable difficulties in Northern Ireland, the application of the Brake must be targeted to those parts. Once the UK Government is convinced that the Brake has been justifiably initiated, it must notify the Commission that it intends to prevent the operation of the EU law in question and make a detailed explanation of how the conditions for the Brake have been fulfilled. At this point the new article 13(3a) process dovetails into the existing article 13(4) process, which applies to the inclusion of new EU goods rules within annex 2, allowing for a six-week window for the exchange of views between the UK Government and the Commission. This can then generate a Joint Committee response if both agree to a means of addressing the issue, or to the Commission taking remedial action if it concludes that the Brake has been misapplied. Uses of the Brake therefore come with attendant risks to Northern Ireland's access to the EU Single Market. Either the decision to invoke the Brake or any EU counter-measures can thereafter be the subject of the Withdrawal Agreement's arbitration arrangements, which would aim to determine if all of the requirements for using the mechanism have been fulfilled.[44]

The new feature works alongside the existing article 13(4) process by which the UK Government can prevent new EU law measures from being added to the scope of the Protocol, giving a new route to achieve this outcome where an annex 2 provision is being amended or replaced. In summary, where there is a significant change to a relevant EU law relating to goods, which is subject to objection under this mechanism, the UK Government will be able to notify the EU and block the operation of this measure in Northern Ireland, or at least those parts which can be demonstrated to be causing 'significant impact' to society in Northern Ireland. This creates divergences in how the two Brake processes apply: article 13(4) covers any new measure within the scope of the Protocol, whereas article 13(3a) covers amendments to existing EU law within the scope of annex 2. If the Brake is misapplied, as with the existing article 13(4) process and the introduction of safeguard measures

44 UK–EU Withdrawal Agreement (n 2 above) art 175.

under article 16,[45] there is scope for arbitration. The new process is, however, more constrained than article 16. The UK Government cannot simply assert some general lack of cross-community support to prevent the amendment or replacement of an aspect of EU law relevant to goods, as it has repeatedly threatened to do in recent years when discussing article 16 (although these suggestions that article 16 has a particularly broad scope have not been formally tested and have been subject to criticism).[46] There is now an expectation that such concerns will be channelled through the Stormont Brake mechanism and meet its procedural requirements.

The Stormont Brake is, as noted above, restricted to EU rules on customs, goods and agriculture, which allows rules applicable to electricity to continue to apply under the Protocol's protections of the all-island Single Electricity Market,[47] as well as leaving the rights and equalities provisions untouched. The limited scope of the Stormont Brake, however, raises interesting questions over how it will apply to complex EU measures; for example, how future EU rules such as the Carbon Border Adjustment Mechanism, which apply to both electricity and goods, will be treated.[48] Such complex and multifaceted measures call into question the bald terms of Sunak's assurance that '[i]f the veto is used, the European courts can never overturn our decision',[49] for there are many questions of EU law which could be left to be determined as part of the Stormont Brake process. Even the concept of 'significant impact' could turn upon conflicting accounts of how an amended EU law will operate, and the CJEU will not relinquish the determination of such questions to an arbitration panel.

---

45 See B Melo Araujo, 'An analysis of the UK Government's defence of the Northern Ireland Protocol Bill under international law' (2022) 73(S2) Northern Ireland Legal Quarterly 89.

46 See R Howse, 'Safeguards' in Federico Fabbrini (ed), *The Law and Politics of Brexit: Volume IV The Protocol on Ireland/Northern Ireland* (Oxford University Press 2022) 253.

47 For more on the Single Electricity Market, see N Robb, 'What does the Windsor Framework mean for the Single Electricity Market?' (*Brexit and Environment* 14 March 2023).

48 The Command Paper accompanying the Framework, moreover, generated an expectation that the same arrangements for Stormont input would apply to the art 13(4) process with regard to new EU law: HM Government (n 21 above) paras 67–68. There was, however, no legal obligation upon the UK Government to extend the mechanism to art 13(4); S Peers, 'Just say no? The new "Stormont Brake" in the Windsor Framework' (*EU Law Analysis* 5 March 2023).

49 Sunak (n 1 above) col 574.

THE CO-OPERATION DIVIDEND

The Protocol Bill, with which the UK Government had threatened to unilaterally disapply parts of the Protocol, has been swiftly waved off the stage. In the UK Government's own legal assessment, 'given the terms of the Windsor Framework and the clear availability of a durable negotiated solution, there would now be no legal justification for enacting the Northern Ireland Protocol Bill'.[50] With it goes the spectre of its underdeveloped dual regulatory arrangements and the uncertainties of its impact on the Protocol's rights and equality arrangements.[51] Perhaps it did its job as a piece of theatre, enabling Sunak, in his visible reluctance to take forward the legislation, to differentiate himself from his predecessor. Johnson's efforts to cast himself as the Conservative Party's once-and-future King have also shaped this deal. The EU has been eager to avoid any return of his bombast and has thus been willing to shore up Sunak with this extensive package of mitigations in the expectation of a better era of relations.

The degree of the shift from the EU's October 2021 proposals is therefore a marker of how much the Commission is willing to invest in this new relationship, and of its willingness to promote the image that technocratic dealing yields results over table thumping, much as Johnson sought to present the Protocol Bill as having 'brought the EU to negotiate seriously'.[52] Even though the ERG and Johnson would eventually announce their opposition to the deal, perhaps in the hope of exploiting any resultant turbulence around its implementation, doing so only served to expose their current parliamentary weakness. The Commission can, moreover, be confident that, given the polls ahead of the looming UK general election, this deal could potentially provide a bridge to further strengthening of cooperation under a Keir Starmer-led Labour Government; one which might see more extensive UK-wide alignment with the EU, and thereby prevent and negate some of the most pressing potential and emerging divergences between Great Britain and Northern Ireland.[53]

Does this mean that the Protocol problem has been solved? When the wheels came off the deal in early 2021, following the eruption over Covid vaccines and the possibility of the EU triggering article 16, it was a scant few days after Boris Johnson had lauded the Joint Committee clarifications on the Protocol's workings. The difference

50 UK Government, 'UK Government legal position: the Windsor Framework' (27 February 2023).

51 See Murray (n 15 above) 28–30.

52 P Walker and L O'Carroll, 'Boris Johnson says he will find it "very difficult" to vote for Northern Ireland deal' *The Guardian* (London 2 March 2023).

53 K Starmer, 'Speech at Queen's University Belfast on the Northern Ireland Protocol' (13 January 2023).

this time round is that the press conference announcing the Windsor Framework was so ebullient; the Sunak Government will struggle to claim that it did not understand the terms it was signing up to, as Boris Johnson attempted to do. This pushes it towards cooperating over the response. Perhaps the most important facet of the Windsor Framework is therefore the accompanying reset in relations. The deal establishes a process for managing the Protocol which will brings the two sides together through extant and new mechanisms which create a dense institutional architecture. Crucially, these structures are inclusive of the Northern Ireland institutions as well as business and civil society.

To the existing institutions, the Joint Committee, Specialised Committee and Joint Consultative Working Group (JCWG), are added a new Special Body on Goods operating as part of the Specialised Committee to provide for exchanges of views on future UK legislation on goods regulation and sub-groups on goods and electricity under the JCWG. Inclusion of both business and civil society responds to requests from those in Northern Ireland and widens the aperture beyond the narrow business-focused structures previously proposed.[54] Crucially, this would allow stakeholders to speak to both the UK and EU at the same time, hopefully bringing a higher quality of understanding to both sides of the negotiation and ensuring that each side is hearing the same testimony. There also remain some outstanding issues, such as on veterinary medicines and the implementation of the Framework which cooperation should facilitate resolution on, as well as on unknown unknowns which arise.

Improved relations should also facilitate movement on other areas of UK–EU cooperation including under the Trade and Cooperation Agreement (TCA). The Commission has already announced that it will begin work on the UK's re-entry to the Horizon Europe programme, a development welcomed by business, charities and universities across the UK and Ireland. Perhaps seeking to save some good news for a future news cycle, the UK Government has, however, been slow to commit itself to the programme.[55] Cooperation in other areas under the TCA which have been stalled, such as in electricity trading and energy security, could be unlocked by more cooperative relationships between the UK and the EU. Sunak's aim must be to present these developments as evidence of the success of his Government's reset in relations with the EU ahead of the next general election.

54 See European Economic and Social Committee, 'The implementation of the EU–UK Withdrawal Agreement, including the Protocol on Ireland and Northern Ireland' (25 January 2023) REX/563-EESC-2022.

55 G Parker, B Staton and A Bounds, 'Rishi Sunak holds back on rejoining Horizon after Brexit breakthrough' *Financial Times* (London 3 March 2023).

IMPLEMENTING THE WINDSOR FRAMEWORK

The nature of the Windsor Framework, being built upon the Protocol's terms and requiring the EU to modify the implementation of EU law in Northern Ireland, negated the need for much active parliamentary engagement at Westminster.[56] The most important implementation activity took place at the EU institutions or in the Joint Committee meeting on 24 March 2023. There have not, after all, been Westminster votes to approve previous Joint Committee decisions, and therefore all that was left for the UK Parliament to vote on was the implementation of the Stormont Brake, which could be achieved using the broad powers to make regulations under the Withdrawal legislation, even where it involves the amendment of primary legislation.[57] In terms of Parliament's involvement in the process, the Stormont Brake became not so much the rabbit out of the hat within the deal, as the whole performance; 'it is a debate and vote on the statutory instrument, but as No 10 has said, that will be taken as an overall say on the Windsor framework itself'.[58] The lack of meaningful parliamentary debate around the deal as a whole, and the publication of the specific Stormont Brake regulations only days before the vote on them, compounded the impression that the UK Government was eager to avoid substantive parliamentary scrutiny of the deal.[59] That the vote on the regulations took place on the same day that Boris Johnson appeared before the Standards Committee with regard to accusations of misleading Parliament over breaches of Covid rules and guidelines, moreover, helped to deflect public attention from the meagre parliamentary opposition the Framework faced. That rebellion, however, when it came, was noteworthy only in how limited it was, serving to underscore this loose coalition's current lack of parliamentary leverage. The Windsor Framework (Democratic Scrutiny) Regulations 2023 passed by 515 votes to 29, with little debate or drama.

These regulations represent an effort to iron out inconsistencies between the UK Government's stated position on the Stormont Brake

56 In addition to the headline-grabbing vote on the Stormont Brake, other regulations have been promulgated which take steps which are now permitted under the deal, including with regard to VAT in Northern Ireland: The Value Added Tax (Installation of Energy-Saving Materials) Order 2023 (SI 2023/376).

57 European Union (Withdrawal) Act 2018, s 8C(1).

58 European Scrutiny Committee (n 27 above) Q116 (Brendan Threlfall, Acting Director General, Cabinet Office).

59 The Lords' Select Committee responsible for reviewing secondary legislation found its scrutiny curtailed by this push to complete the parliamentary process: Secondary Legislation Scrutiny Committee, *34th Report of Session 2022–23* (HL 2023) para 14.

and the legal terms of the Windsor Framework deal, by inserting a new schedule 6B into the Northern Ireland Act 1998.[60] The new schedule pulls off something of a sleight of hand, merging the Stormont arrangements applicable to the (new) articles 13(3a) and (existing) article 13(4) of the Protocol and introducing into both the cross-community voting requirements which were missing from the UK–EU agreement on the Stormont Brake, but were prominent within the UK Government's Command Paper on the deal. The addition of this requirement into the article 13(4) process, for the adoption of new EU law which is within the scope of the Protocol, is not particularly controversial. This amounts to the UK Government adding a new trigger for its absolute power under article 13(4) of the original Protocol to object to the application of new EU law to Northern Ireland. The same cannot, however, be said of the extension of this process to article 13(3a), regarding the modification of existing annex 2 measures. These Stormont Brake processes were carefully agreed by the UK and the EU, and the EU might well have objected on good-faith implementation grounds to this belated addition of a process which makes the Brake rather closer to a Unionist veto than the terms of article 13(3a) allow.[61] Nonetheless, the UK Government will have calculated that the EU will not want to rock the boat on a deal which promises to curtail the melodrama surrounding the Protocol. The cross-community vote, moreover, comes late in the Stormont Brake process, and the EU will have some confidence that it can troubleshoot problems in Committee ahead of the Brake being employed, warn UK ministers off its use by publishing potential retaliatory measures or challenge any misuse through arbitration. The domestic functioning of the provisions of the Protocol remains a matter for the UK Government, provided that it continues to meet its treaty obligations, including good-faith implementation.

These arrangements also foresee the possibility of judicial review challenges to the UK Government's approach to the Brake and are presented as efforts to tie the Government into following up on objections from the Assembly, asserting in particular that 'the possibility of the European Union taking remedial measures in accordance with Article 13(4) of the Framework is not a relevant consideration'.[62] This, of course, is easier to assert than to establish, for the potential of a triggering of the Brake to cause difficulties for UK–EU relations will never be fully excluded from ministerial thinking, even if it does not make it into the (required) public reasons for a UK

60 Windsor Framework (Democratic Scrutiny) Regulations 2023 (n 43 above).
61 UK–EU Withdrawal Agreement (n 2 above) art 5.
62 Northern Ireland Act 1998, sch 6B, para 14(2).

Government decision not to employ the Brake.[63] It remains open for ministers to say that they do not believe that the broadly framed 'significant impact' conditions for using the Brake have been met. The UK Government is seeking to present these arrangements to Unionists as imposing considerable strictures upon ministerial discretion,[64] thereby making it appear that the Government will be required to follow Stormont's lead if sufficient MLAs object, but the courts can be expected to accord ministers considerable latitude in such decision-making.

The resultant Stormont Brake arrangements are not, as their critics would portray them, 'useless in practice'.[65] It is not designed for general operation, but to respond to particular measures where Northern Ireland's representatives can demonstrate that significant issues on the ground in Northern Ireland have not been properly taken into account in the development of EU law. Notwithstanding the Brake's complexity, it operates alongside the existing article 18 Stormont Lock to give the Northern Ireland Assembly a meaningful say over both the overarching nature of the trade arrangements applicable to Northern Ireland under the Withdrawal Agreement and the implementation of specific EU law which adds to or amends those arrangements. It operates in a way which builds upon the Petition of Concern processes instantiated in the 1998 Agreement, and subject to subsequent refinement,[66] and it is this close attention to this special feature of Northern Ireland's governance order which generates much of the complexity in its terms.

The expectation is that the Brake becomes a channel for Unionist concerns over the development of EU law, and thereby dissuade the Unionist parties, if power-sharing is functioning, from reaching for other tools like the 'St Andrews veto', which involves Northern Ireland Ministers seeking the Executive Committee's approval for 'significant or controversial matters', and could thereby be employed to prevent them from proceeding with aspects of EU law which require transposition.[67] The Northern Ireland High Court has looked askance at efforts by the DUP's Agriculture Minister Edwin Poots to use this mechanism to attempt to block the commencement of port infrastructure as required by the Protocol,[68] but the Windsor Framework offers a preferable

63 Ibid para 16.

64 See European Scrutiny Committee (n 27 above) Q87 (Mark Davies, Director, Windsor Framework Taskforce).

65 M Howe and B Reynolds, 'The European Research Group's Legal Advisory Committee Review and Assessment of the "The Windsor Framework"' (22 March 2023) para 8.

66 Northern Ireland (Ministers, Elections and Petitions of Concern) Act 2022, s 6.

67 Northern Ireland Act 1998, s 20(4).

68 *In re Rooney* [2022] NIKB 34, [198] (Colton J).

alternative to such obstructionist tools, which in this instance ultimately saw the UK Parliament legislate in Stormont's stead.[69] These options nonetheless remain on the table if power-sharing is restored, and with many Unionist representatives unreconciled to Northern Ireland's post-Brexit trade arrangements, they might be employed in concert to stymie those arrangements. Should that occur, much will depend on how indulgent the UK Government is towards such efforts, given that Westminster has overlapping powers to implement the Protocol as amended by the Windsor Framework,[70] against the risk of appropriate 'remedial action' by the EU.

THE BEST OF BOTH WORLDS: PART II?

The Windsor Framework mitigates some of the most pressing difficulties with the operation of Protocol's trade rules, and the arrangements put in place thus more plausibly support claims that Northern Ireland-based businesses will be able to benefit from the access they will enjoy to the EU Single Market (in terms of goods) and the UK internal market. But it does so in a way which generates a much more complex picture of the applicable legal rules, and one which could be subject to rival interpretations. As the EU's post-Brexit relationship with the UK continues to develop, the arrangements will, moreover, require careful ongoing management against the backdrop of the heightened fragility of Northern Ireland's constitutional politics; there will undoubtedly be other difficulties down the line which will need to be navigated. If the power-sharing institutions are restored, there will come a point when the 'Stormont Brake' will be on the table and, perhaps sooner, when the EU identifies a particular risk to the Single Market and makes efforts to have checks expanded. These looming flashpoints mean that the world's 'most exciting economic zone'[71] remains, in some ways, potentially too exciting.

The deal does not, moreover, mitigate all of the challenges of potential divergence in product standards for manufactured goods between Great Britain and Northern Ireland, particularly if the Retained EU Law (Revocation and Reform) Bill is enacted in its current form. For the UK Government, any post-Brexit divergences in these product standards will generate potential challenges for Northern Ireland, and it might

69 The Official Controls (Northern Ireland) Regulations 2023 (SI 2023/17).

70 The UK Government cannot neglect its duties to implement the Withdrawal Agreement as a result of difficulties securing implementation in the Northern Ireland Assembly: Vienna Convention on the Law of Treaties 1969 (1980) 1155 UNTS 331, art 27.

71 A Forrest, 'Rishi Sunak mocked for calling Northern Ireland "world's most exciting economic zone"' *The Independent* (London 28 February 2023).

become more predictable for law in Great Britain to track many EU developments applicable to Northern Ireland than attempt to cajole Unionist parties in the Assembly into using the Stormont Brake.[72] In contrast to the scope for post-Brexit divergence from EU law in Great Britain, every time the Stormont Brake is used to prevent developments in EU law relating to goods taking effect in Northern Ireland law, the pre-existing EU law will continue to apply. Goods rules operative in Northern Ireland which look neither like the rules in place in Great Britain nor the EU would likely become very difficult for businesses to manage.[73] That practical issue might ultimately constrain the Brake's use, as much as its convoluted procedural requirements.

Notwithstanding those shortcomings, progress under the Windsor Framework comes in the form of process. Many of the changes will be phased in over the next two years, giving businesses an extended period to adapt to the new requirements (in a marked change from the short grace periods provided in early 2021, when businesses were still struggling to unpack the relationship between the Protocol and the TCA). Moreover, the UK Government and EU appear to have jointly accepted the invidious consequences of using Northern Ireland as leverage in redrawing their post-Brexit relationship and to be committed to managing challenges through the Withdrawal Agreement's technocratic committee systems. Northern Ireland's stability rests on these issues being kept out of the headlines.

---

72 The significance of international standards for exporters (the impact of the 'Brussels effect') and the UK's TCA commitments regarding these standards also impacts upon the potential for divergence with regard to manufactured goods: HM Government (n 21 above) para 58.

73 This is analogous to the application of non-diminution rules under art 2 of the Protocol, which requires Northern Ireland law to continue to apply EU law as it existed at the point in time the UK withdrew from the EU; see C Murray and C Rice, 'Beyond trade: implementing the Ireland/Northern Ireland Protocol's human rights and equalities provisions' (2021) 72 Northern Ireland Legal Quarterly 1, 28.

Articles

Commentaries and Notes