The City Law Review
Volume IV
2022

The Editorial Board
Volume IV
2022

The City Law Review

Volume IV

2022

Proudly endorsed by the City Law School

The City Law Review (the 'CLR') is City, University of London's student-managed, peer-reviewed, publication of legal scholarship. This year, we are proud to have been sponsored by Bryan Cave Leighton Paisner LLP, who have been the earliest supporters of the CLR and have continued to generously sponsor Volume IV as well as an award in their name for Best Diverse Piece. Thank you to Mishcon de Reya LLP who have graciously sponsored this year's addition as well as the Writer's Excellence Award in their name and No5 Chambers for sponsoring the publication of this volume. We are primarily funded by the City Law School, without the endorsement of which we would not function.

The objective of the CLR is to provide a space whereby students can have their work published, and the Editorial Board can be exposed to legal writing and responsibility not often afforded to undergraduates. The CLR operates through a double-blind peer-review process, organised by our Editorial Board, meaning all pieces retain anonymity until the final draft at which point, they undergo academic review by selected City, University of London staff. Prizes have been awarded in the same fashion, voted on by the Editorial Board based on merit alone. The categories are Writer's Award for Excellence, Best Diversity Piece, and Women's Recognition in Law Award. Under the guidance of Shabana Elshazly, Sophia Evans, and Jonathan Lynch, the review has afforded legal academics a forum to produce and discuss legal concepts and current debates. This process, originally set up by former Editor-in-Chief, Shabana Elshazly, ensures the robustness of the CLR and has provided an invaluable foundation upon which this Volume, and future Volumes, are built.

This year, we looked at ways in which we could further establish the highly regarded reputation of the CLR and we are delighted that all submissions are available on City, University of London's Online Research Database (CRO) as well as the British Library. This allows for the online publication of pieces onto a research database to be used, and referenced, by students for years to come. We hope that this will allow future Volumes to continue growing, reputation and reach.

The views expressed by the contributors are not necessarily those of the CLR, the Editorial Board, the City Law School or our sponsors. This publication is intended to be a conduit for the scholarship of the student body. While every effort has been made to correct and develop the articles, the accuracy and completeness of information is the duty of each author individually. The CLR does not assume responsibility for any factual errors, misquotations, misleading representations, or inconsistencies.

The CLR uses the Fourth Edition of Oxford University Standard for Citation of Legal Authorities ('OSCOLA').

British Library Cataloguing in Publication for Information
A catalogue record for this book is available at the British Library
ISSN: 2516-2160

Published in the United Kingdom
By the City Law Review
City, University of London
Northampton Square
London, EC1V

TABLE OF CONTENTS

Editor's Note

We are proud to present the fourth Volume of the *City Law Review*.

The articles we have published during my term as Editor-in-Chief and, indeed, over the four-year history of the *City Law Review*, include exciting contributions to legal academia. Volume IV contains articles with envelope-pushing ideas on everything from defining transgender parenthood to the efficacy of the international criminal court. The *City Law Review* was first established in 2013 and was formally titled the *City Law Society Journal.* Since its rebranding, each editorial board has worked tirelessly to grow and evolve the *Review* into what it is today. As City's first and sole student-led publication of legal scholarship, we are focused on providing an evolving platform to ensure the longevity of future Volumes. As my predecessors have state, it is an honour to contribute to this growing legacy.

The articles in Volume IV present new and innovative approaches to divergent legal topics and denote new lenses through which to examine the law, touching upon a variety of areas of interest. This year, we are proud to be sponsored by Bryan Cave Leighton Paisner LLP, Mishcon de Reya LLP and No5 Chambers. We are delighted to have an in-person launch event in March 2022 with the panel including Siân Mirchandani QC, Gemma McNeil Walsh, and Christianah Babajide who will partake in a discussion on the topic of 'Women in Law.' Thank you to our guest speakers. We are grateful for the time and effort you took to share your thoughts and experiences.

The Review particularly expresses gratitude to the City Law School and Executive Committee who have been pillars of support in the production of each Volume. In particular, Dean Professor Andrew Stockley, for his commitment to ensuring the success of Volume IV and promoting the *Review's* recognition within the law school. Special thanks are owed to Dr David Seymour, The *Review's* longest-standing supporter, for his unparalleled support throughout my three years on the Editorial Board. Not only is he tasked with the responsibility of the organisation and operation of our Academic Review process, which is crucial in ensuring the integrity and legal accuracy of our submissions, but he has been a guiding light in the creation of the Review. David has supported my ideas, inspired me to carry out my action plans, and has always provided the most thoughtful advice. He has made my time as Editor-in-Chief a truly remarkable experience. The support of the faculty members of the City Law School by offering guidance and direction on student scholarship is a testament to their dedication to legal academia and further reinforces the foundation upon which each Volume is built.

The origin of the *Review* stems back to Shabiir Bokhari, who spearheaded the *Review* in its early days. He has continued to be a key individual in shaping the growth of the volume by offering significant advice while seeking supporters and encouragement from key institutions. Recognition must be made to Sophia Evans who had trusted me as the sole first-year Article Editor. I have had the opportunity to be taken under the wing of two past Editor-in-Chief's: Shabana El Shazly and Jonathan Lynch. Shabana was instrumental in the rebranding of the Review and curating our Patron System. She has guided me throughout my time in this position, speaking words of encouragement and inspiration. Jonathan Lynch and Emily Woolf have not only been my mentors but prepared me to undertake this position. Their dedication and passion for the Review are qualities that I admire most about both of them. I will forever be grateful for Jonathan's unwavering faith and confidence in my potential to be the Editor-in-Chief.

Production of this *Review* involved the talent and effort of our esteemed editorial board. A very special thank you to those who served as Article Editors and Senior Article Editors. Priya Ahsan Chowdhury's dedication to her role as Senior Article Editor was evident throughout this year by always providing significant insight while working from the other side of the world. Daniel Hale Bolingbroke and Ishani Thakrar have exceeded expectations. They were always the first Editor's to offer support and assistance when needed and demonstrated a considerably high standard of work ethic. Daniel, your eagerness and commitment to the Review will suit you well as the forthcoming Editor-in-Chief.

To the *Review's* Publishing Editor Eryn Green, there are no words to express the gratitude I feel for your work on this special volume. You have been instrumental from the initial consideration of this project through its printing. I could not be prouder of the work that you have contributed throughout the year. To our Deputy-in-Chief Monica Kiosseva, you do it all and the Volume IV team have been immensely lucky to benefit from your leadership. Monica has overlooked the work of our two Managing Editors, Soraya Arif and Saba *Tabassum* who have also been crucial in the administrative operation of Volume IV. Responsible for co-managing all writer correspondence, Soraya and Saba have worked hard to ensure the smooth running of all departments. It has been rewarding to motivate and inspire a team of intelligent and excited individuals who have worked so collaboratively to create something so remarkable. They have been a truly outstanding team, and I am grateful for having the opportunity to lead them.

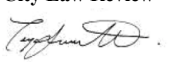
It has been an immense honour to serve as the Editor-in-Chief of *The City Law Review* and guide the publication of Volume IV this forthcoming year. I have been fortunate to watch the progression of the Law Review since my initial contribution as an Article Editor three years ago. To finalise this year's volume on the eve of graduation has been a wonderful reminder of the importance of legal scholarship in influencing litigation, policy, and public discourse. I am privileged to encourage law students to contribute to the editorial process and management of the publication as this process is an active form of learning. Our goal is to shape legal dialogue through the specific pieces we select for publication and thus the consequentiality of the ideas we put forward. I am excited to have built an intellectual community within City that focuses on legal scholarship and prominent contemporary dialogue.

I would once again like to thank the City Law School for their continued support. Lastly, I would like to thank my family for always motivating me to strive to accomplish my goals. They have raised me to become the focused and ambitious woman I am today and have been my pillars through this position as well as in life.

I hope you enjoy the carefully crafted collection as well as the important critiques of the legal landscape surfacing in each area of law included in Volume IV.

Yours Faithfully,

Teya Fiorante
Editor-in-Chief
City Law Review



Foreword

I am delighted to have been invited to write a foreword to this issue of the City Law Review.

Like many of the very best American law journals, the City Law Review is student-led and student-edited. It provides a wonderful opportunity for some of our best students to be involved in all aspects of producing a law journal, from soliciting and reading a wide range of work, deciding what should be published, developing all the skills needed to edit legal writing, and overseeing the publication, marketing, and distribution processes. I congratulate the Editor-in-Chief, Jonathon Lynch, and all the other students involved in producing this volume. They have done especially well given the circumstances of the last year and all the restrictions of the Covid-19 pandemic. The cover picture shows the new City Law School building in Sebastian Street completed towards the end of 2020. As at the time of writing this foreword, we have not yet been able to occupy it but look forward to doing so later this year!

This is a journal that aims to publish some of the very best of our students' research. One of the strengths of the City Law School is that we teach law at all levels, from apprenticeships to the LLB, LLM and PhD degrees, from the Graduate Diploma in Law for graduates of other disciplines, to the Bar Vocational Studies Course and the Legal Practice Course for intending barristers and solicitors. Having formerly been the Inns of Court School of Law we have a proud and distinguished history of legal education. Students from all parts of the School have an opportunity to submit work for the City Law Review and this volume shows the variety of legal and topical issues some of them have been researching and writing on.

My congratulations to everyone involved for their enthusiasm and hard work. The editors and contributors can be very proud of this issue of the City Law Review.

Professor Andrew Stockley

Dean of The City Law School

The Intrusion Lacuna and Tort Law's Corrective Justice Purpose

Nazanin Taher

Introduction

The current protection afforded to the right of privacy within tort law fails to protect against pure, physical intrusions. As such, this essay argues that the law must be reformed to protect against such intrusions. It will first discuss tort's purpose of corrective justice and recognise that the intrusion lacuna that currently persists in English tort doctrine prevents this purpose from being fulfilled. This intrusion lacuna will then be compared to the recent global developments in introducing intrusion-type torts, further highlighting the need to modernise the doctrine. This essay will then propose the development of an 'intrusion upon seclusion' tort and address the main critiques against this development. Ultimately, it will conclude that introducing an 'intrusion upon seclusion' tort is a necessary and legitimate reform that will bring coherence to the doctrine by fulfilling tort law's corrective justice purpose.

The Purpose of Tort Law

The purpose of tort law is corrective justice, which 'imposes on wrongdoers the duty to repair their wrongs and the wrongful losses on their wrongdoing occasions.'[1] Specifically, tort law is best informed by Perry's volitional approach where the 'focus [is] on the normative implications of voluntary action.'[2] This is perhaps most obvious in *Donoghue,*[3] where a commitment to corrective justice '[justified] engaging in quite radical development of the common law.'[4] In the judgement of *Donoghue*, Lord Atkin suggested it would be wrong 'to deny a legal remedy where there is so obviously a social wrong.'[5] In *X (minors),*[6] Lord Bingham stated that corrective justice has the 'first claim on the loyalty of the law.'[7] Indeed, alternative forms of justice lead to the 'risk of leaving individuals without relief for substantial unjust losses,'[8] as Coleman argues. These judgements make it clear that corrective justice 'supports fair and coherent determinations of liability,'[9] which not only is but ought to be the purpose of tort law.

Summary of the Doctrine

[1] Jules L Coleman, 'The Mixed Conception of Corrective Justice' [1992] 77 Iowa L Rev 427, 441.
[2] Stephen R Perry, 'The Moral Foundations of Tort Law' [1992] 77 Iowa L Rev 449, 451.
[3] *Donoghue v Stevenson* [1932] AC 562 (HL).
[4] Thomas DC Bennett, 'Privacy, Corrective Justice and Incrementalism: Legal Imagination and the Recognition of a Privacy Tort in Ontario' [2013] 59 McGill L J 49, 72.
[5] *Donoghue v Stevenson* [1932] AC 562 (HL), [583] (Lord Atkin).
[6] *X (Minors) v Bedforeshire CC* [1995] 2 AC 633 (HL).
[7] ibid, [663] (Lord Bingham).
[8] Jules L Coleman, 'Tort Law and the Demands of Corrective Justice' [1992] 67 Ind LJ 349, 359.
[9] Ernest Weinrib, *Corrective Justice* (1st edn, OUP 2012) 7.

The current privacy doctrine needs reform as the intrusion lacuna prevents the fulfilment of corrective justice. Under English jurisdiction, 'no specific privacy tort exists.'[10] However, some protection of informational privacy is afforded through the 'misuse of private information' (MPI) doctrine, established in *Campbell*[11] and confirmed to be a tort in *Vidal-Hall.*[12] MPI developed from the equitable doctrine of confidence and followed a 'new methodology.'[13] This allowed information that was no longer confidential to be protected, as in *PJS,*[14] thus resolving a deficiency of the confidence doctrine. The development of MPI was a response to the incorporation of the European Convention on Human Rights (ECHR) into domestic law through the Human Rights Act 1998 (HRA),[15] which provides the right to 'private and family life' under Article 8.[16] The ECHR has been given a form of indirect horizontal effect. Therefore, the courts must render the law to be compatible under s.6 HRA[17] and develop mechanisms for individuals to bring claims against rights violations.

However, as the MPI doctrine only protects private information, physical intrusions remain unprotected. In *Kaye*[18], where two journalists intruded upon the claimant's hospital room, the Court of Appeal reaffirmed that there was no actionable right of privacy in English law and, as such, could not rectify the claimant on that basis. In this case, Lord Bingham stated that 'this invasion of his privacy…however gross, does not entitle him to relief in English law.'[19] In *Wainwright,*[20] it was held that the claimants who were subjected to humiliating and distressing strip search carried out 'sloppily' during a prison visit were unable to seek any remedy despite suffering from emotional distress and PTSD as a result. Whilst the court could remedy *Kaye* through the doctrine of malicious falsehood, as the journalists had attained information through photographs and an interview, *Wainwright* failed entirely as 'there was no question of there being any private information (or photographs) being disclosed.'[21]. The intrusion lacuna is not only unjust but is also unlawful according to the ECtHR, which found violations of Article 8 and 13 of the ECHR[22] in *Wainwright.*[23] This portrays a failure of the current doctrine in fulfilling tort law's purpose of corrective justice because of the absence of remedies for the kind of harm caused to the claimants in the case of *Wainwright*. Ultimately, this is incoherent with the rest of tort doctrine which does correct wrongdoing and must be subject to reform.

[10] Kirsty Horsey & Erika Rackley, *Tort Law* (6th edn, OUP 2019) 447.
[11] *Campbell v MGN Ltd* [2004] UKHL 22; [2004] AC 457.
[12] *Vidal-Hall v Google Inc* [2015] EWCA Civ 311.
[13] Rebecca Moosavian, 'Charting the journey from confidence to the new methodology' [2012] EIPR 34(5), 324.
[14] *PJS v NGN Ltd* [2016] UKSC 26, [2016] AC 1081.
[15] Thomas DC Bennett, 'Privacy, Corrective Justice and Incrementalism: Legal Imagination and the Recognition of a Privacy Tort in Ontario' [2013] 59 McGill L J 49, 54.
[16] Convention for the Protection of Human Rights and Fundamental Freedoms (European Convention on Human Rights, as amended) (ECHR), art 8.
[17] Human Rights Act 1998, s 6.
[18] *Kaye v Robertson* [1990] FSR 62 (CA).
[19] ibid, [70] (Lord Bingham).
[20] *Wainwright v Home Office* [2004] 2 AC 406 (HL).
[21] Kirsty Horsey & Erika Rackley, *Tort Law* (6th edn, OUP 2019) 478.
[22] Convention for the Protection of Human Rights and Fundamental Freedoms (European Convention on Human Rights, as amended) (ECHR), art 13.
[23] *Wainwright v UK* [2008] 1 PLR 398.

Comparison to Foreign Jurisdictions

There has been a global shift in recognising torts dealing with intrusion, particularly in other common law countries and in comparison, England currently finds itself behind.[24] The United States has had four distinct privacy torts for around half a century,[25] including an 'intrusion upon seclusion' tort, established from the work of Prosser.[26] In the case of *Holland,*[27] concerning a man 'who surreptitiously videoed his flatmate in the shower',[28] a judgement passed in the high court of New Zealand led to the development of a new 'intrusion into seclusion' tort. In Australia, the high court has 'expressly' left the door open for recognising a common law right to privacy,[29] and the Queensland District Court recognised the tort of intrusion in the judgement of *Grosse.*[30]

Interestingly, Ontarian law, which had remarkably similar privacy laws to that of England, recognised an 'intrusion upon seclusion' tort in the case of *Jones.*[31] Here, the act of the defendant merely accessing the plaintiff's banking records was held to be a violation of privacy despite the information not having been published, distributed or recorded. In giving the lead judgement, Sharpe JA first established that whilst 'the courts did not accept the existence of a privacy tort, they rarely went so far as to rule out the potential of such a tort.'[32] He also considered the 'explicit recognition of a right to privacy'[33] as a Canadian Charter value.[34] As Canadian Charter Values are given indirect horizontal effect,[35] Sharpe JA suggested that 'common law should be developed in a manner consistent with [such values].'[36] He further justifies this development through Craig's 'principled approach,'[37] which suggests 'courts… [may] create new categories of torts …[to] give effect to overarching principle.'[38] Shape JA implies corrective justice is this overarching principle through his focus on the causal relationship between the wrongful 'deliberate, prolonged, and shocking'[39] conduct of the defendant and causing harm to the plaintiff who 'would be profoundly disturbed by the…intrusion.'[40] Through this, he legitimises the recognition of an 'intrusion upon seclusion'

[24] Thomas DC Bennett, 'Privacy, Corrective Justice and Incrementalism: Legal Imagination and the Recognition of a Privacy Tort in Ontario' [2013] 59 McGill L J 49, 56.
[25] Restatement (Second) of Torts [1977], s 652.
[26] William L Prosser, 'Privacy' [1960] 48(3) Cal LR 383.
[27] *C v Holland* [2012] NZHC 2155.
[28] N A Moreham, 'Liability for listening: why phone hacking is an actionable breach of privacy' [2015] 7(2) Journal of Media Law 155, 163.
[29] Thomas DC Bennett, 'Privacy, Corrective Justice and Incrementalism: Legal Imagination and the Recognition of a Privacy Tort in Ontario' [2013] 59 McGill L J 49, 56.
[30] *Grosse v Purvis* [2003] QDC 151.
[31] *Jones v Tsige* [2012] ONCA 32.
[32] ibid, [31] (Sharpe JA).
[33] ibid, [45] (Sharpe JA).
[34] Canadian Charter of Rights and Freedoms, Part 1 of the Constitution Act, 1982, being Schedule B to the Canada Act 1982 (UK), c 11; *Hunter v Southam Inc* [1984] 2 SCR 145.
[35] *Hill v Church of Scientology* [1995] 2 SCR 1130.
[36] *Jones v Tsige* [2012] ONCA 32, [46] (Sharpe JA).
[37] John D R Craig, 'Invasion of Privacy and Charter Values: The Common-Law Tort Awakens' [1997] 42 McGill L J 355.
[38] Thomas DC Bennett, 'Privacy, Corrective Justice and Incrementalism: Legal Imagination and the Recognition of a Privacy Tort in Ontario' [2013] 59 McGill L J 49, 82.
[39] *Jones v Tsige* [2012] ONCA 32, [69] (Sharpe JA).
[40] ibid.

tort, as this 'would amount to an incremental step that is consistent with the role of this court to develop the common law.'[41]

A similar development to that in *Jones* is possible within English privacy doctrine as English courts have similarly not ruled out the possibility of an 'intrusion upon seclusion' tort. Indeed, *Tchenguiz*[42] 'contains statements seriously suggesting that MPI might provide a remedy in circumstances where private information has been wrongfully acquired but not published.'[43] Indirect horizontal effect is also given to the ECHR, which comparably recognises the right to privacy. As such, our common law should develop compatibility. Furthermore, English courts may exercise wide incrementalism,[44] which like the principled approach, allows for novel causes of action through regard to overarching principles, such as corrective justice. Indeed, Sharpe JA's 'view that Ontario law would be 'sadly deficient' if it did not provide Jones with a remedy mirrors Lord Atkins concern that failing to recognise a general duty of care would be a 'grave defect in the law… so contrary to principle.'[45] Therefore, the English courts can and should modernise the doctrine to protect against intrusions in this way.

Proposed Reform

It is proposed that the courts should develop a new 'intrusion upon seclusion' tort, based on the reasoning in *Jones*, which aims to correct '[physical intrusions] into a private space where someone is in a state of seclusion'[46] and 'the use of the senses to oversee or overhear a person's private affairs.'[47] This reform also aims to modernise the doctrine in light of the global shift towards recognising torts of intrusion and render it compatible with Article 8 of the ECHR.

The elements of this tort of intrusion, based on elements in *Holland* and *Jones*, should be as follows: '(a) an intentional and unauthorised intrusion; (b) into seclusion (namely intimate personal activity, space or affairs); (c) involving infringement of a reasonable expectation of privacy; (d) That is highly offensive to a reasonable person;'[48] (e) which has caused 'distress, humiliation or anguish.'[49] As corrective justice 'necessitates a strong focus on harm'[50] and causation, these elements focus on the cause, type and severity of the harm suffered. The reasonable expectation of privacy requirement has been included as this reflects the language of the ECtHR jurisprudence, thus ensuring greater conformity with the ECHR. Despite not receiving support in *Campbell,*[51] the 'highly offensive' requirement has been included to

[41] *Jones v Tsige* [2012] ONCA 32, [65] (Sharpe JA).
[42] *Tchenguiz v Imerman* [2010] EWCA Civ 908, [2011] Fam 116.
[43] Thomas DC Bennett, 'Privacy, Corrective Justice and Incrementalism: Legal Imagination and the Recognition of a Privacy Tort in Ontario' [2013] 59 McGill L J 49, 56.
[44] Lesley Dolding and Richard Mullender, 'Tort Law, Incrementalism, and the House of Lords' [1996] 47 N Ir Legal Q 12.
[45] Thomas DC Bennett, 'Privacy, Corrective Justice and Incrementalism: Legal Imagination and the Recognition of a Privacy Tort in Ontario' [2013] 59 McGill L J 49, 74.
[46] John Hartshorne, 'The Need for an Intrusion upon Seclusion Privacy Tort within English Law' [2017] 46 Comm L World Rev 287, 289.
[47] ibid.
[48] *C v Holland* [2012] NZHC 2155, [2012] 3 NZLR 672, [94].
[49] *Jones v Tsige* [2012] ONCA 32, [71].
[50] Thomas DC Bennett, 'Privacy, Corrective Justice and Incrementalism: Legal Imagination and the Recognition of a Privacy Tort in Ontario' [2013] 59 McGill L J 49, 72.
[51] Paul Wragg, 'Recognising a Privacy-Invasion Tort: The Conceptual Unity of Informational and Intrusion Claims' [2019] CLJ 78(2), 414.

balance concerns of 'floodgates of litigation'[52] and the 'importance of free speech, which often conflicts with privacy claims.'[53] In this way, this reform will serve corrective justice while balancing consequentialist concerns.

Defence of Proposed Reform

There are three main critiques of developing an 'intrusion upon seclusion' tort. The first is that the introduction of a new cause of action amounts to 'judicial activism' and is beyond the powers of the courts. Formalists such as Ewing suggest that 'it is not the job of the judicial branch to make the law.'[54] Lady Hale also suggests 'the courts will not invent a new cause of action to cover types of activity which were not previously covered.'[55] However, the courts may legitimately exercise wide incrementalism,[56] 'having regard to overarching principles… to find novel causes of action.'[57] This is seen in both *Donoghue*[58] and *Jones,*[59] where the judges were concerned about dealing with the causal relationship between the defendant and claimant for providing a remedy even though there was no tightly analogous precedent confirming the existence of a cause of action. Furthermore, Tugendhat J contends that 'courts may legitimately extend… the law quite significantly.'[60] Therefore, developing an 'intrusion upon seclusion' tort based on the overarching principle of corrective justice would amount to an exercise of wide incrementalism and would be within the 'constitutional constraints'[61] of the courts.

The second critique is that the right to privacy under Article 8 ECHR is already adequately protected. In *Wainwright*, Lord Hoffman argued against finding 'a breach of Article 8'[62] as this would 'only demonstrate that there was a gap in the English remedies for invasion of privacy which has since been filled by section 6 and 7 of the [HRA].'[63] Lord Hoffman, like Posner,[64] appears to have conceptualised privacy solely as control over information. However, this conceptualisation fails to 'take individuals seriously'[65] and is inconsistent with the view of the ECtHR, which has confirmed that physical violations of a person's physical or

[52] Thomas DC Bennett, 'Privacy, Corrective Justice and Incrementalism: Legal Imagination and the Recognition of a Privacy Tort in Ontario' [2013] 59 McGill L J 49, 66.
[53] Thomas DC Bennett, 'Privacy, Corrective Justice and Incrementalism: Legal Imagination and the Recognition of a Privacy Tort in Ontario' [2013] 59 McGill L J 49, 67.
[54] KD Ewing, 'A Theory of Democratic Adjudication: Towards a Representative, Accountable and Independent Judiciary' [2003] 38(3) Alberta LR 708, 710.
[55] *Campbell v MGN Ltd* [2004] UKHL 22, [2004] AC 457, [495] (Baroness Hale).
[56] Lesley Dolding and Richard Mullender, 'Tort Law, Incrementalism, and the House of Lords' [1996] 47 N Ir Legal Q 12.
[57] Thomas DC Bennett, 'Privacy, Corrective Justice and Incrementalism: Legal Imagination and the Recognition of a Privacy Tort in Ontario' [2013] 59 McGill L J 49, 83.
[58] *Donoghue v Stevenson* [1932] AC 562 (HL).
[59] *Jones v Tsige* [2012] ONCA 32.
[60] Thomas DC Bennett, 'Judicial activism and the nature of 'misuse of private information'' [2018] 23(2) Comms L 74, 85.
[61] Gavin Phillipson and Alexander Williams, 'Horizontal Effect and the Constitutional Constraint' [2011] 74(6) MLR 878.
[62] *Wainwright v Home Office* [2004] 2 AC 406 (HL), [52] (Lord Hoffmann).
[63] ibid.
[64] Richard A Posner, 'Privacy, Secrecy, and Reputation' [1979] 28 Buffalo Law Review 1.
[65] Allan C Hutchinson and Derek Morgan, 'The Canengusian Connection: The Kaleidoscope of Tort Theory' [1984] 22 Osgoode Hall L J 69, 87.

psychological integrity will amount to a breach of the convention. Indeed, in *X and Y*[66] and *Söderman,*[67] the domestic legislation's lack of remedy or deterrent criminal measure to guard against physical intrusions was held to breach Article 8 of the ECHR. As such, the court enforced a positive obligation on Sweden to put in place legal provisions that would avail victims in those situations. As the courts must consider ECHR jurisprudence under s.2 of the HRA[68] and are currently at risk of a similar positive obligation imposed from the ECtHR, it is clear that a tort of 'intrusion upon seclusion' is necessary to enable compatibility with the ECHR and adequately protect the right to privacy.

The third criticism is that English courts already recognise the tort of intrusion. Moreham suggests that *Gulati,*[69] a case involving phone hacking, 'should be seen as a part of this trend towards greater protection for physical privacy interests.'[70] Hartshorne agrees and suggests *Gulati* allows 'individuals to pursue an MPI claim before the English courts in… circumstances where in both Ontario and New Zealand… [gave rise to] the recognition of an 'intrusion upon seclusion' tort.'[71] However, Rowbottom points out that '[t]he decision in *Gulati* was only to assess damages and the scope of the tort was not the issue before the court.'[72] Additionally, it 'rendered exclusively under the doctrine of MPI'[73] and did not deal with physical intrusions. Thus, it cannot be concluded that the English courts recognise the tort of intrusion.

Conclusion

This essay has argued that English tort doctrine requires reform and has proposed the development of an 'intrusion upon seclusion' tort. It is established that the purpose of tort law is corrective justice. Thus, such reform will bring greater coherence within the doctrine through remedying violations of physical intrusions which cannot be rectified under the current doctrine. Contrary to formalist criticisms, this development is legitimate through exercise of wide incrementalism by the English courts. Moreover, the stated reform is necessary for adequate protection of Article 8 of the ECHR, which recognises the protection of physical integrity and to modernise the doctrine of English tort law in line with global developments.

[66] *X and Y v The Netherlands* (1985) 8 EHRR 235.
[67] *Söderman v Sweden* [2014] 58 EHRR 36.
[68] Human Rights Act 1998, s 2.
[69] *Gulati and others v MGN Ltd* [2015] EWCA Civ 129, [2015] WLR(D) 232.
[70] N A Moreham, 'Liability for listening: why phone hacking is an actionable breach of privacy' [2015] 7(2) Journal of Media Law 155, 164.
[71] John Hartshorne, 'The Need for an Intrusion upon Seclusion Privacy Tort within English Law' [2017] 46 Comm L World Rev 287, 297.
[72] Jacob Rowbottom, 'A landmark at a turning point: Campbell and the use of privacy law to constrain media power' [2015] 7(2) Journal of Media Law 170, 186.
[73] Thomas DC Bennett, 'Triangulating Intrusion in Privacy Law' [2019] 39(4) OJLS 751, 752.

Strict Product Liability Law and New Technologies: The Heavy Burden to Prove a Robot Caused Injury

Emelia Kovivisto

Abstract

As advanced technologies integrate seamlessly and profoundly into our societies, matters of dispute grow increasingly complex. Machine-related accidents are harder to explain, and it becomes a more challenging endeavour to show how they caused personal injury. An illustration of the looming intricacy of evidence questions in complex injury litigation has recently surfaced in the U.S. in the field of medical robotics. While widely used, highly advanced, and soon coupled with artificial intelligence technologies, robotic surgical tools have repeatedly entailed defects for which no liability was found. This article explores how U.S. courts regarded evidence of machine malfunction and causation in two recent cases of robot-assisted surgeries. It examines how both plaintiffs unsuccessfully attempted to, first, provide direct evidence of a product defect, and second, proffer circumstantial evidence that the defect was indeed the proximate cause for their postoperative injuries. The article discusses the high evidentiary standards imposed by the two courts against the backdrop of similar incidences in recent years. Using the example of medical robotics, this article seeks to illuminate how the burden of proof under current liability laws risks becoming excessively heavy, if not insurmountable, for future plaintiffs injured by complex technologies.

Introduction

Evidence is filtered fact. Ideally, fault-based litigation applies a purifying filter to determine who ought to rightfully bear responsibility for a harm caused. As advanced technologies integrate seamlessly and profoundly into our societies, matters of dispute grow increasingly complex. How a machine-related accident caused a given personal injury becomes increasingly challenging to prove before a court of law. Novel questions of evidential adequacy, sufficiency and reliability will soon be posed in intricate cases where new technologies collide with traditional product liability law.

One of the existing and perhaps most illustrative areas of rising complexity in product litigation has surfaced in the United States in the field of surgical robotics. In the past decades, a notable number of cases have concerned the da Vinci surgical system, a long-armed robot operating on patients under the remote command of the surgeon.[1] Several instances of machine defects have given rise to product liability claims against da Vinci's manufacturer, Intuitive Surgical Inc.[2] Thousands of lawsuits have been filed against the company, most of which were

[1] United States Securities and Exchange Commission, 'Intuitive Surgical, Inc. Form (10-Q) Q1 Quarterly Report Pursuant to Section 13 or 15(d) of the Securities Exchange Act of 1934' (2021) 22 <https://docoh.com/filing/1035267/0001035267-21-000067/ISRG-10Q-2021Q1> accessed 15 May 2021.

[2] Data compiled by the Food and Drug Administration (FDA) suggests that the device induced 144 deaths and 1391 injuries between 2000 and 2013 (respectively 1,4% and 13% of all interventions conducted). Homa

settled behind closed doors.[3] Of the few cases that have reached a courtroom, the majority resulted in successful defences by Intuitive, and not one has entailed a full verdict for the plaintiff.[4]

Two cases among those so far litigated are exemplary for this study.[5] Both plaintiffs underwent a da Vinci-assisted surgery and experienced severe postoperative complications. Neither of them was able to prove that the robot-assisted surgery caused their injuries, and both were denied summary judgement as a result. Unlike other lawsuits against Intuitive that concerned surgeon credentialing,[6] medical malpractice[7] or procedural questions,[8] the two chosen cases centred around the admissibility and sufficiency of evidence. For this research the cases are particularly illustrative of evidentiary hurdles in complex product litigation as they exemplify how two types of evidence – direct and circumstantial – were deemed by different federal courts as insufficient. The factual congruence, the similarities in the parties' arguments and the identical evidentiary questions in both cases serve as a common denominator creating a basis for exploring the boundaries of strict product liability in the field of advanced technologies.

This article examines how courts have viewed evidence of defect and causation in accidents connected to robot-assisted surgeries. The hypothesis guiding this research puts forth that, given the advent of increasingly intelligent technologies, a strong evidentiary filter may curtail future plaintiff's abilities to bring personal injury claims. The first part of the article presents the facts of the two cases and the courts' reasons for denying summary judgement to the plaintiffs. The second part focuses on direct evidence by exploring the legal framework, the plaintiffs' proofs, and reasons why neither was able to demonstrate a product defect. The third part turns to the element of circumstantial evidence and examines the legal context, the arguments of the plaintiffs, and the courts' standards of causal evidence which both failed to meet. The conclusion reflects on ways in which high standards for proof risk to impede future strict liability action.

This article's protagonist is the da Vinci surgical system, a contemporary example of high technological sophistication. Using computational, robotic, and advanced imaging

Alemzadeh and others, 'Adverse Events in Robotic Surgery: A Retrospective Study of 14 Years of FDA Data' [2016] 11 PLOS ONE e0151470.

[3] A recent class action involving an alleged 2000 patients was settled by Intuitive for $42.5 million in California. ibid; Docket Entry, *In re Intuitive Surgical Securities Litigation*, No 5:13-cv-1920 (ND Cal, Sept 11, 2018).

[4] It must be said that a few plaintiffs were not left completely empty-handed. In *Mendoza v Intuitive*, the court upheld the plaintiff's claims in part, as the case concerned a previously recalled model of da Vinci. The case *Taylor v Intuitive* was challenged up to the third instance which, for procedural errors, vacated the verdict and remanded the case for retrial without a final ruling. *Mendoza v Intuitive Surgical, Inc*., No. 18-CV-06414-LHK, 2020 US Dist LEXIS 73780 (ND Cal Apr 24, 2020); *Taylor v Intuitive Surgical, Inc.,* 187 Wn.2d 743 (Wash 2017); cf *O'Brien v Intuitive Surgical, Inc*., No. 10 C 3005, 2011 US Dist LEXIS 80868 (ND Ill Jul 25, 2011) (dismissing the plaintiff's complaint against Intuitive as unsubstantiated).

[5] *Mracek v Bryn Mawr Hosp*., 610 F Supp 2d 401 (ED Pa 2009) and *Pierre v Intuitive Surgical, Inc*., 476 F Supp 3d 1260 (SD Fla 2020).

[6] *Mohler v St Luke's Med Ctr*., No 1 CA-CV 08-0078, 2008 Ariz App LEXIS 492 (Ariz Ct App Dec 26, 2008) (surrounding the question whether the surgeon had proper authorization and training to use the robotic device).

[7] *Dulski v Intuitive Surgical, Inc*., No 10-CV-0234A, 2011 US Dist LEXIS 12651 (WDNY, Jan 19, 2011) (regarding a medical malpractice action against the treating surgeons).

[8] *Taylor v Intuitive Surgical, Inc.,* 187 Wn.2d 743 (Wash 2017) (concerning the adequacy of jury instructions about the manufacturer's failure to warn about risks associated with the da Vinci system).

technologies, the device translates the hand motions of a surgeon working at a console into precise movements inside the patient.[9] The high level of precision and dexterity allows for less invasive interventions compared to conventional procedures.[10] The da Vinci system obtained clearance from the U.S. Food and Drug Administration (FDA) in 2000 as the first and only device approved for various minimally invasive surgeries, a position it held for over a decade.[11] It continues to be widely employed for a variety of minimally invasive surgeries, most commonly for urologic and gynaecologic interventions.[12] Such treatment was what the plaintiffs Roland Mracek and Elmitha Pierre also sought and received, ultimately to their own misfortune.

The Cases of Mracek and Pierre

The case of Roland Mracek

In October 2004, Roland Mracek underwent a prostate biopsy at Bryn Mawr Hospital, Pennsylvania and was diagnosed with adenocarcinoma, a type of prostate cancer.[13] After receiving the diagnosis, he consulted with his urologist Dr McGinnis about treatment options and expressed his concern about developing postoperative complications.[14] As a result, Dr McGinnis recommended a radical prostatectomy[15] using the da Vinci system.[16] The use of the robot would allow the intervention to be less invasive than a traditional laparoscopic surgery and would minimise the risk of complications.[17]

Mracek consented to the suggested treatment, and the surgery commenced on June 9th, 2005.[18] In the midst of the intervention, the robot suddenly stopped functioning and displayed red error messages.[19] Dr McGinnis' surgical team attempted to restart the robot several times, and a representative of Intuitive was called to troubleshoot the machine.[20] Despite all efforts,

[9] 'Intuitive Surgical, Inc. (Form 10-Q) Q1 Quarterly Report' (n 1) 22.
[10] Alemzadeh and others (n 2) 1.
[11] The da Vinci was the only FDA-approved surgical robot authorised for urologic, gynaecologic, cardiac, and head and neck procedures for fifteen years after its commercialization. ibid; see also Jonathan Douissard, Monika Hagen and P Morel, 'The Da Vinci Surgical System' in Carlos Eduardo Domene and others (eds.), *Bariatric Robotic Surgery* (Springer 2019) 8.
[12] ibid.
[13] *Mracek v Bryn Mawr Hosp*., 610 F Supp 2d 401 (E.D. Pa. 2009). Adenocarcinoma is a type of cancer that forms in glandular cells of the body, such as the prostate gland. Maurie Markman, 'Adenocarcinoma - Types and Treatment Options' (*Cancer Treatment Centers of America*, 1 November 2018) <https://www.cancercenter.com/adenocarcinoma> accessed 23 March 2021.
[14] *Mracek,* 610 F Supp 2d at 402.
[15] A robot-assisted radical prostatectomy is conducted by making several small keyhole incisions in the abdomen of the patient through which the prostate is then removed. The intervention is less invasive than a conventional laparoscopic surgery which entails a cut from the patient's belly button down to the pubic bone. See 'Robotic Prostatectomy' (*Johns Hopkins Medicine*, 2021) <https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/robotic-prostatectomy> accessed 4 May 2021.
[16] Brief for Appellant, *Mracek v Bryn Mawr Hosp*., 363 Fed Appx 925, at 3 (3d Cir. 2010) (stating that Mracek consented to the surgery relying on the recommendation of Dr McGinnis) (as cited in Christopher Beglinger, 'Note: A Broken Theory: The Malfunction Theory of Strict Products Liability and the Need for a New Doctrine in the Field of Surgical Robotics' [2019] 104 Minnesota Law Review 1041, 1042).
[17] *Mracek,* 610 F Supp 2d at 402.
[18] ibid.
[19] ibid., [403].
[20] ibid.

the robot failed to function. The surgical team was ultimately required to shift to a conventional surgery and finished the intervention through a traditional laparoscopic procedure.[21] A week after the surgery, Mracek suffered gross haematuria[22] and was readmitted for further hospitalisation.[23] He contended to suffer long-term consequences of the operation such as total erectile dysfunction and severe daily abdominal pain.[24] In 2008, Mracek raised a claim against Intuitive in strict product liability and strict malfunction liability. [25]

To establish the product liability claim, Mracek argued that the defect was obvious.[26] He relied on the fact that the robot repeatedly flashed red signs of 'error', that it shut down and that it could not be restarted to finish the surgery.[27] He argued that the defect was sufficiently apparent to be ascertainable to the jury and contended that he did not need to substantiate the evidence with expert testimony at the pre-trial stage.[28] The Court of the Eastern District of Pennsylvania disagreed and held that the surgical robot was too complex for the jury to infer a defect.[29] It stated that the average juror could not reach a conclusion without speculation in the absence of any expert testimony.[30] The court held that Mracek's lack of direct evidence of a defect was fatal to his strict product liability claim.[31]

As another cause of action, Mracek claimed strict malfunction liability. Under the malfunction theory in Pennsylvania, the plaintiff may prove a product defect by presenting circumstantial evidence that a malfunction occurred, that the product was not used in an abnormal way and that no other causes accounted for the harm.[32] The court held that Mracek failed to eliminate all other reasonable secondary causes, as the complexity of the medical procedure allowed several other reasons to contribute to his injuries.[33] In short, Mracek failed to show sufficient causal salience between the defect and his injury. The court thus held that Mracek had failed to bear his burden of circumstantial evidence to support his strict malfunction liability claim.[34] As a result, the court denied Mracek the motion for summary judgement on all claims which was affirmed on appeal a year later.[35]

The case of Elmitha Pierre

[21] ibid.
[22] Gross hematuria is defined as the visible occurrence of blood in a person's urine. Jeanne Charleston, 'Hematuria (Blood in the Urine)' (*National Institute of Diabetes and Digestive and Kidney Diseases*, 2021) <https://www.niddk.nih.gov/health-information/urologic-diseases/hematuria-blood-urine> accessed 4 May 2021.
[23] *Mracek,* 610 F Supp 2d at 403.
[24] ibid.
[25] ibid; Mracek's other claims of breach of warranty and negligence were deemed to fail for the same evidence reasons as his actions under strict product liability and strict malfunction liability. The claims were not discussed in detail by the courts and are thus immaterial to this analysis.
[26] *Mracek,* 610 F Supp 2d at 405.
[27] ibid.
[28] ibid.; Mracek did not contest that expert testimony was required after pre-trial in the summary judgement hearing.
[29] ibid.
[30] ibid., [405-406].
[31] ibid., [407].
[32] ibid., [408].
[33] ibid.
[34] ibid.
[35] ibid., [409]; *Mracek v Bryn Mawr Hosp*., 363 Fed Appx 925 (3d Cir. 2010).

A similar fate dawned on Elmitha Pierre in Florida.[36] Various types of treatment had proven unsuccessful for treating her symptomatic uterine fibroid.[37] Her treating physician Dr Chen then proposed a surgical removal of her uterus.[38] Following his recommendation that a da Vinci-assisted laparoscopic hysterectomy[39] was less invasive and post-operative risks were lower than in a traditional hysterectomy, Pierre underwent the surgery on January 24th, 2014.[40] Close to the end of the operation, Dr Chen noticed that Pierre's bowel had been damaged.[41] He consulted with the hospital's general surgeon who examined the damaged tissue and concluded that it was not severe enough to require additional surgical intervention.[42] Pierre was kept in the hospital for observation, and no additional repairs were made to her bowel.[43]

Six days after the surgery, Pierre experienced severe physical side effects that she alleged were caused during the surgery.[44] Pierre specifically contended that the electrosurgical scissors at the end of one of da Vinci's robotic arms had a broken insulation through which thermal energy had leaked and caused burns, or electric arcing to her bowel.[45] She filed a motion for summary judgement against Intuitive before the Southern District Court of Florida in 2018, claiming strict product liability.[46]

Florida courts generally apply the consumer expectation test on strict product liability claims.[47] Pierre argued that the instrument frustrated consumers' safety expectations given that the scissors must have been defective since arcing had occurred during the surgery.[48] The court contested Pierre's argument and asserted that she could not demonstrate the occurrence of arcing without admissible expert testimony.[49] Despite Dr Chen having spotted damage to her tissue during the operation, the court held that there was, 'no direct evidence indicating that the

[36] *Pierre v Intuitive Surgical, Inc*., 476 F Supp 3d 1260 (SD Fla 2020).
[37] ibid 1266; Uterine fibroids are benign pelvic tumours. Surgical removal is recommended as a curative solution in cases where the tumour is particularly symptomatic. Geum Seon Sohn and others, 'Current Medical Treatment of Uterine Fibroids' [2018] 61 Obstetrics & Gynecology Science 192, 192; 'Fibroids: What Are Fibroids? Fibroids Symptoms, Treatment, Diagnosis' (*UCLA Health*) <https://www.uclahealth.org/fibroids/what-are-fibroids> accessed 6 April 2021.
[38] ibid.
[39] ibid.; Hysterectomy is the term for a surgical removal of the uterus. A conventional hysterectomy is conducted by an open incision, typically six to twelve inches long, through the abdominal wall of the patient. In a robot-assisted laparoscopic hysterectomy, multiple small incisions are made in the patient's abdomen making the intervention less invasive and the recovery time shorter. Melissa McMacken, 'Laparoscopic Hysterectomy' (*Johns Hopkins Medicine*, 2020) <https://www.hopkinsmedicine.org/gynecology_obstetrics/specialty_areas/gynecological_services/treatments_services/minimally_invasive_gynecologic_robotic_surgery/treatments/laparoscopic_hysterectomy.html> accessed 6 April 2021; see also Jon I Einarsson and Yoko Suzuki, 'Total Laparoscopic Hysterectomy: 10 Steps Toward a Successful Procedure' [2009] 2 Reviews in Obstetrics and Gynecology 57, 57.
[40] ibid.
[41] ibid.
[42] ibid.
[43] ibid.
[44] ibid., [1267].
[45] ibid., [1265-1266].
[46] Pierre also claimed a breach of duty of care, failure to warn and loss of consortium with her husband, none of which prevailed for summary judgement. Examination of these claims lies outside the scope of this article.
[47] The Court also applied the risk utility and the reasonable alternative design test to assess a product defect which survived. Ultimately however, Pierre's claims of defect all failed because her evidence of causation was deemed insufficient.
[48] *Pierre*, 476 F Supp 3d, [1271].
[49] ibid.

[i]nstrument did not perform as expected during Pierre's surgery.'[50] The court thus held that her strict liability claim did not prevail under the consumer expectation test.

To establish causation between the alleged defect and her injury, Pierre introduced expert testimony as circumstantial evidence.[51] Specifically, Pierre relied on Dr Chen's expert opinion that the likely cause for Pierre's injuries was arcing caused by a leak of thermal energy.[52] Dr Chen provided evidence of causation through a differential diagnosis[53] opening on the most probable cause. However, the court held that his statement contained insufficient certainty to present a reliable piece of causal evidence.[54] The court concluded that Dr Chen's expert evidence was inadmissible, and that Pierre had thus failed to prove causation.[55] Pierre was denied summary judgement in March 2020, which was affirmed on appeal in April 2021.[56]

Both Mracek and Pierre thereby failed to meet their burden of proof to survive a motion for summary judgement. The next two sections turn to examine more closely how the plaintiffs' direct evidence of product defect and their circumstantial evidence of causation were insufficient to claim strict product liability.

Direct Evidence as First Type of Proof

The current legal framework of U.S. strict product liability

The emergence of modern strict product liability law was generated by the 20th century societal transformation. The post-industrial shift to mass production, to developed supply chains and to efficient means of transportation created physical distance between the sellers of products and their buyers.[57] These factors limited consumers' abilities to hold manufacturers accountable for defective products under the old law.[58] To meet the demand for enhanced consumer safety, product liability law shifted from placing responsibility on the buyer (caveat emptor) to imposing it on the seller (caveat venditor).[59] The idea became that manufacturers ought to be strictly responsible for harm occurring through the use of their products, as they were in a better position to mitigate risks.[60]

[50] ibid.
[51] Pierre introduced testimony provided by several experts, all of which were rejected by the court as either conjectural evidence or hearsay; ibid., [1267].
[52] ibid., [1267].
[53] Differential diagnosis as a process of providing evidence through which 'a physician systematically eliminates possible causes of a patient's ailment to arrive at its most probable cause', ibid 1275.
[54] ibid., [1277].
[55] ibid.
[56] *Pierre v Intuitive Surgical, Inc.,* No 20-11311, 2021 US App LEXIS 9623 (11th Cir Apr 2, 2021).
[57] Kyle Graham, 'Strict Products Liability at 50: Four Histories' [2014] 98 Marquette Law Review 555, 561.
[58] Typical barriers to bringing a claim were requirements that the manufacturer and the consumer had to be in contractual privity (the privity rule), difficulties to prove the producer's negligence or express warranties that commonly excluded coverage of personal injury damage. ibid 567; George L Priest, ''Strict Products Liability: The Original Intent' by George L. Priest' [1989] 10 Cordozo Law Review 2301, 2305; Chad E Wallace and Andrew T Wampler, 'Skimming the Trout from the Milk: Using Circumstantial Evidence to Prove Product Defects under the Restatement (Third) of Torts: Products Liability Section 3, Tennessee and Beyond' [2000] 68 Tennessee Law Review 647, 650.
[59] Graham (n 57), [568]; Priest (n 58), [2305].
[60] ibid.

Since the 1960s, § 402A of Restatement (Second) of Torts has governed strict product liability law in the United States.[61] Its adoption was seen as revolutionising liability in tort as it expanded the scope of liability, as well as stringently separated allocation of responsibility from considerations of a defendant's possible care.[62] Today, most states in the U.S. have adopted § 402A as governing law of strict product liability,[63] including Pennsylvania[64] and Florida.[65] Importantly and despite the prevailing notion of caveat venditor, it remains on the plaintiff to show that product defect existed and caused the harm.[66] As the plaintiff is seen as the party who seeks the relief, § 402A provides that the burden of proof is to be carried by the plaintiff.[67]

How Mracek and Pierre failed to proffer direct evidence

Following § 402A, the Pennsylvania Eastern District Court required Mracek to show that the product was defective, that the defect existed at the time the product was in the manufacturer's control and that the defect proximately caused his injuries.[68] The court specified that a plaintiff does not need to support direct evidence with expert testimony, if the matter is 'simple […] and within the range of comprehension of the average juror.'[69] Mracek claimed that this was the case.[70] Given da Vinci's repeated flashing of 'error' messages and its total shutdown during the surgery, Mracek contended that inferring a defect was not beyond a lay person's apprehension, and that he therefore did not need expert testimony to establish a defect.[71]

The court disagreed with Mracek.[72] Instead, it held that the da Vinci system was too complex of a machine to permit the jury to find it dysfunctional without the assistance of an expert.[73] The court agreed with Intuitive that the average juror would not have the requisite background to reach a conclusion devoid of speculation.[74] Without expert testimony, Mracek's

[61] David Owen, 'Defectiveness Restated: Exploding the 'Strict' Products Liability Myth' [1996] 1996 University of Illinois Law Review 743, 747–784; Beglinger (n 16) 1051.
[62] It laid down the dividing line between strict liability and negligence. Aaron Twerski and others, 'The Technological Expert in Products Liability Litigation' [1974] 52 Texas Law Review 1303, 1303; Priest (n 58) 2301; William Prosser, 'The Assault upon the Citadel (Strict Liability to the Consumer)' [1960] 69 Yale Law Journal 1009, 1112.
[63] John CP Goldberg and Benjamin C Zipursky, 'The Strict Liability in Fault and the Fault in Strict Liability' [2016] 85 Fordham Law Review 743, 745.
[64] § 402A was adopted in Pennsylvania in *Webb v Zern*, where a beer manufacturer was strictly liable for injuries caused by an exploding beer keg. *Webb v Zern*, 422 Pa 424 (Pa 1966); see also Mansman Jerome J., 'Torts - Products Liability - Restatement (Second), Torts, 402(A) - The Pennsylvania Supreme Court Adopts a Strict Tort Liability Rule for the Products Liability Area.' [1966] 5 Duquesne Law Review 215, 217.
[65] Florida adopted § 402A in the judgement of *West v Caterpillar Tractor Co*. as leading authority on strict liability. *West v Caterpillar Tractor Co*., 336 So 2d 80 (Fla 1976); see also Spencer H Silverglate, 'The Restatement (Third) of Torts Products Liability: The Tension Between Product Design and Product Warnings' [2001] 75 The Florida Bar Journal 10, 10.
[66] George S Jr Mahaffey, 'Cause-in-Fact and the Plaintiff's Burden of Proof with Regard to Causation and Damages in Transactional Legal Malpractice Matters: The Necessity of Demonstrating the Better Deal' [2004] 37 Suffolk University Law Review 393, 407.
[67] *Restatement (Second) of Torts* § 402A cmt g (Am Law Inst 1965).
[68] *Mracek*, 610 F Supp 2d, [404].
[69] ibid., [405].
[70] ibid.
[71] ibid.
[72] ibid.
[73] ibid.
[74] ibid., [406].

strict product liability claim had to fail. On appeal, the Third Circuit Court affirmed this reasoning.[75] Despite reiterating that no expert testimony was legally necessary where other evidence allowed the jury to 'clearly see' the construction and use of the machine, it ruled that this was not the case for Mracek.[76] The court implied that the factual record of the incident did not enable the jury to clearly see a defect.[77] Mracek's direct evidence was confirmed insufficient, as the product was held to be too complex for the jury.

The complexity of the da Vinci posed a similar hurdle for Pierre. In Florida, courts have followed § 402A in strict product liability cases by customarily applying the consumer expectation test. The test determines a product to be defective when it fails to perform as safely as the ordinary consumer would expect.[78] Florida law stipulates that the consumer expectation test effectively reflects the policy considerations of strict product liability, as it imputes responsibility on the basis that the product's marketing and selling create reasonable expectations of its performance to the consumer which a defect would naturally frustrate.[79] Nevertheless, courts in Florida have acknowledged, in accordance with other states' courts,[80] that the test is inadequate for distinctively complex products.[81]

Given the complexity of the da Vinci, the first question for the Southern District Court of Florida was whether the test would be applicable to Pierre's case at all. It regarded Intuitive's argument as not meritless in maintaining that the product was too complex for the ordinary consumer to have reasonable expectations about its proper operation.[82] It also remarked that da Vinci-like medical devices were generally marketed to hospitals and medical professionals and not to ordinary consumers.[83] Nevertheless, the court was more persuaded by the strong policy considerations prevalent in Florida law and emphasised that the consumer expectation test 'best vindicates the purposes underlying the doctrine of strict liability.'[84] As a result, the court did

[75] *Mracek*, 363 F App 925, 927 (3d Cir. 2010).
[76] ibid.
[77] For a detailed discussion about the appellate court's reasoning on this particular issue, see Beglinger (n 16) 1076.
[78] The basis for the consumer expectation test is found in comment g to § 402A: 'Where the product is, at the time it leaves the seller's hands, in a condition not contemplated by the ultimate consumer, which will be unreasonably dangerous to him.' Restatement (Second) of Torts § 402A cmt. g. (Am. Law Inst. 1965); see also Tiffany Colt, 'The Resurrection of the Consumer Expectations Test: Regression in American Products Liability' [2018] 26 University of Miami International and Comparative Law Review 525, 543.
[79] ibid 546. *West v Caterpillar* established that 'The manufacturer, by placing on the market a potentially dangerous product for use and consumption and by inducement and promotion encouraging the use of these products, thereby undertakes a certain and special responsibility toward the consuming public who may be injured by it.' (*West v Caterpillar*, 336 So 2d at 86).
[80] For example, *Ray by Holman v BIC Corp*., 925 SW 2d 527, 531 (Tenn 1996) (the Supreme Court of Tennessee deemed the test to be inapplicable as the complexity of a cigarette lighter precluded consumers to hold safety expectations); *Camacho v Honda Motor Co.*, 741 P 2d 1240, 1246-1247 (Colo 1987) (the Colorado Supreme Court held that the test was inadequate to determine complex product liability cases involving technical and scientific products); *Soule v General Motors Corp*., 882 P 2d 298, 308 (Cali 1994) (the California Supreme Court reiterated that the consumer expectation test is not appropriate in cases where the ordinary consumer would have 'no idea how [the product] would perform in all foreseeable situations').
[81] *Cassisi v Maytag Co.,* 396 So 2d 1140, 1145 (Fla Dist Ct App. 1981) (stating that the test is a vague and unreliable complex product litigation because ordinary consumers cannot be assumed to have safety expectations for complex products).
[82] *Pierre*, 476 F Supp 3d, [1270].
[83] ibid., [1271].
[84] ibid.

not depart from the consumer expectation test and required Pierre to show that the product did not perform as safely as an ordinary consumer would expect when it was used as intended.[85]

Pierre argued that the scissors failed to perform along her expectations since electric arcing had occurred.[86] The court contested her argument stating that she had not demonstrated that arcing had in fact eventuated during the surgery.[87] The mere instance of a bowel injury did not suffice. Dr Chen's observations were not admissible in this regard since he did not witness arcing as such, but only assumed arcing to have occurred to explain the damaged tissue.[88] The court held that she had not proffered any direct evidence upon which the lay jury could reasonably rely on to infer a product defect.[89] On appeal, the Eleventh Circuit Court affirmed the judgement and highlighted that the matters at hand 'presented complex medical and scientific issues outside the scope of a layperson's knowledge, so expert testimony was required.' [90] Thus, Pierre's claim failed as the jury was deemed inapt to adjudicate on whether the product's performance frustrated the expectations of an ordinary consumer. Like for Mracek, the complexity of the da Vinci rendered the matter beyond the purview of the common man.

When is a product defect too complex for the common man?

Where the limits to jury's competencies lie have been abundantly debated among legal scholars.[91] Yet, the jury's ostensible inability *per se* was not the reason why neither Mracek's nor Pierre's claims survived. The issue was that the courts deemed the da Vinci system too complex for the jury to examine without expert assistance. By not granting summary judgement to the plaintiffs, both courts precluded the juries from reviewing the cases presuming that the matter was too complex for common people to infer a defect.

In several other strict liability cases concerning complex industrial products, courts have similarly held that plaintiffs could not successfully provide direct evidence without expert testimony. Such cases have occurred in the context of an escalator,[92] a front bumper of a truck,[93]

[85] ibid.
[86] ibid., [1271].
[87] ibid.
[88] ibid. Dr Chen's assumption also relied heavily on inadmissible hearsay of another doctor at the hospital, deceased at the time of the trial.
[89] ibid.
[90] *Pierre v Intuitive Surgical, Inc.,* No. 20-11311, 2021 US App LEXIS 9623, 11 (11th Cir Apr 2, 2021).
[91] For example, L Perrin, 'Expert Witness Testimony: Back to the Future' [1995] 29 University of Richmond Law Review 1389, 1404; Neil Vidmar, 'Are Juries Competent to Decide Liability in Tort Cases Involving Scientific/Medical Issues? Some Data From Medical Malpractice' [1994] 43 Emory Law Journal, Duke Law Scholarship 885.
[92] *Esturban v Mass. Bay Transp. Auth*., 68 Mass App Ct 911, 911 (Mass App Ct 2007) (holding that without expert testimony the plaintiffs failed in showing that there was a greater probability than not that an escalator accident occurred on the account of the defendant as 'an escalator is a complex technical piece of machinery' and 'beyond the scope of an average person's knowledge').
[93] *Oddi v Ford Motor Co*., 234 F.3d 136, 159 (3d Cir 2000) (affirming the district court's judgement for the truck manufacturer and holding that the jury could not infer a product defect by examining a front bumper and flooring of a truck that injured the plaintiff, because 'such conclusions are within the peculiar competence of experts').

a car frame,[94] an airbag,[95] and even something as seemingly simple as a cigarette lighter.[96] In these cases, expert evidence was necessary to substantiate the claim of a product defect.

In other cases, however, comparable products were deemed not too complex for the jury. These cases concerned injuries resulting from a conveyor belt,[97] a van's cargo[98] or a rotating chicken cutter.[99] While the courts did not pronounce these matters to be simple or the proof of defect to be free from uncertainty, there was enough evidence that the courts 'could not exclude the possibility that plaintiff's non-expert evidence will be sufficient to submit his claim of defect to the jury.'[100] In these cases the courts let the cases proceed to trial without expert opinion for there was a *possibility* that the evidence was reasonably apprehensible to the jury.

Similar possibilities were arguably discernible in both *Mracek* and in *Pierre*. The factual record provided by Mracek fathomably encompassed a comprehensible dysfunction. While the product itself is hardly of a simple type, the nature of its defect – the repeated red error messages and the shutdown – does not conceivably exclude the possibility that the defect was in the purview of a layperson. Similarly, the court in *Pierre* did not depart from the consumer expectation test, a yardstick that is generally deemed inadequate for distinctively complex matters.[101] As the court did not depart from the test, it must have assumed that ordinary consumers could hold justifiable expectations about the robot's performance. It logically implied that the machine could not operate entirely beyond the comprehension of the lay juror, who ultimately presents the average consumer. In other words, the fact alone that the court applied the consumer expectation test means that the jury could potentially ascertain a defect within the machine. Yet, it remains difficult to opine on how Mracek and Pierre could have persuaded the courts that evidence of a mere possibility for the jury to grasp the matter would be sufficient.

Notwithstanding the fog of uncertainty surrounding the boundaries of a lay juror's apprehension, courts will expectedly regard future technological devices to be at the more complex end of things (the sophistication of robotic technologies arguably surpasses that of cigarette lighters). At the same time and in contrast to Mracek and Pierre, future plaintiffs might

[94] *Show v Ford Motor Co*., 697 F Supp 2d 975, 985-981 (ND Ill 2010) (holding that an accident induced by a Ford Explorer rolling over was too complex for the jury on the basis that 'unlike a case involving a simple product or even a straightforward brake failure, the jury cannot be expected to understand the dynamics of vehicle stability based on their common experience').
[95] *Wheeler v Chrysler Corp.*, No. 98 C 3875, 2000 U.S. Dist. LEXIS 2725, 10 (ND Ill Feb 29, 2000) (stating that the airbag of an automobile is 'beyond the realm of common understanding' and too complex to be ascertainable by the jury).
[96] *Ray by Holman v BIC Corp*., 925 SW 2d 527, 531 (Tenn 1996) (holding that the complexity of the product precluded ordinary consumers from having reasonable expectations).
[97] *Aldridge v Reckart Equip. Co.,* 2006 Ohio 4964, 52-53 (Ohio Ct App 2006) (stating that an injury arising from a debarking machine's conveyer was not too complex for the jury to adjudicate on defect).
[98] *Atkins v GMC*, 132 Ohio App 3d 556, 564 (Ohio Ct App 1999) (finding that expert testimony was not necessary to prove a defect in a van's cargo, as 'the product nor its allegedly defective aspect [were] so complex as to require expert testimony as a matter of law').
[99] *Padillas v Stork-Gamco, Inc.,* 186 F.3d 412, 416 (3d Cir 1999) (holding that the exposed blades of a rotating chicken cutter did not require expert evidence and that 'it is premature to rule out that [non-expert] testimony and pictures may enable the jury to clearly to see the construction of the machine').
[100] ibid., [416].
[101] n 79.

be even less able to access evidence where defects, or indications of such, go unnoticed during the use of the product. In *Mracek*, the defect was indisputably visible to the bystanders by the fact that it forestalled the whole surgery, and in *Pierre*, the operating surgeon noticed the damaged tissue immediately.[102] It is fathomable that future patients might not be as 'lucky' that incidence will be spotted instantly. Instead, they may be left to grapple with only the long-term repercussions of the accidents which they would then need to trace back to a product defect when bringing a claim before a court. In these cases, the complexity of evidence will rise further and, in turn, render the possibility even slimmer that the jury would be apt to find a defect. In the context of future technologies, high standards of proof risk imposing a heavy burden of direct evidence to show product defects. The question arises then what options of direct evidence plaintiffs are left with, or whether expert testimony will become a de facto imperative for future product liability claims. This stirs concerns about liability law shifting away from the policy idea of strict fault allocation in the name of consumer protection.

Ultimately, it would be incorrect to assume that the law is plainly oblivious to possible unattainability of evidence. Courts have developed alternative ways for plaintiffs to show proof in cases where direct evidence is not accessible. The next section will turn to examine how Mracek and Pierre both attempted to use such alternative routes of evidence and what came in their way.

Circumstantial Evidence as Second Type of Proof

The legal context of circumstantial evidence in strict product liability

While courts have consistently held that the burden to prove a product defect is to be carried by the plaintiff, courts have also increasingly acknowledged that plaintiffs may be unable to point to a specific defect for reasons of destroyed evidence[103] or increased remoteness of the plaintiff from the product.[104] As a result, courts have adopted principles to allow plaintiffs to provide indirect or circumstantial evidence. In those cases, juries are instructed to find that a defect caused the harm on the sole basis that such harm would not occur within an ordinary sequence of events without defect.[105]

Construing cause from circumstantial evidence stems from the doctrine of res ipsa loquitur, meaning 'the thing speaks for itself.'[106] The old rule of evidence, coined in a 19th century English negligence case,[107] allows imputing liability without specific proof of how the harm was caused, but by the mere occurrence of an unusual accident without which the harm

[102] *Mracek*, 610 F Supp 2d at 402; *Pierre*, 476 F Supp 3d, [1267].
[103] David Owen, 'Manufacturing Defects' [2001] 53 South Carolina Law Review 851, 859.
[104] Beglinger (n 16), 1053.
[105] Wallace and Wampler (n 58), 649.
[106] 'Res Ipsa Loquitur' <https://thelawdictionary.org/article/res-ipsa-things-speak/> accessed 2 March 2021; For a more detailed discussion on the relationship between circumstantial evidence and the doctrine of res ipsa loquitur, see William Prosser, 'Handbook of the Law of Torts' [1941] 4 Louisiana Law Review 156, 162.
[107] *Byrne v Boadle*, 159 ER 299 (Exch 1963).

would not have resulted.[108] The maxim is generally viewed as an 'evidentiary short cut'[109] permitting inferences of causation in the absence of first-hand evidence. Most states have developed principles derived from the res ipsa doctrine to allow circumstantial evidence to establish causation in strict liability.[110] Two of these principles are the malfunction theory and the differential diagnosis of an expert, as were respectively used by Mracek and Pierre.

How Mracek and Pierre failed to provide circumstantial evidence

Mracek brought a malfunction liability claim under the Pennsylvanian circumstantial evidence doctrine, the malfunction theory.[111] Being a res ipsa rule, the doctrine allows the plaintiff to prove a product defect without specifically demonstrating how it was defective.[112] For a malfunction liability claim to prevail under the theory, the plaintiff must prove that a malfunction occurred, that the product was used as intended and that no secondary causes contributed to the harm.[113]

The court held that Mracek failed to eliminate other reasonable causes for his injuries.[114] It relied on Intuitive's assertion that several secondary causes could have accounted for the harm, given the intricacy of the robot-assisted surgery and the 'use and timing of various ancillary medical equipment in connection with this innovative and complex procedure.'[115] In addition, the court stated that Mracek's operating physician had not pointed to any definitive causal connection between the robot's malfunction and Mracek's post-operative trauma that would reasonably exclude other secondary causes.[116] The appellate court affirmed the ruling and added that Mracek's circumstantial evidence could not establish malfunction liability as it did not allow that 'a rational finder of fact could find in his favo[u]r.'[117]

[108] Restatement (Second) of Torts § 328D. See also W. PROSSER, LAW OF TORTS § 39 (4th ed. 1971), cited in 'Circumstantial Evidence in Strict Products Liability Actions, *Lindsay v McDonnell Douglas Aircraft Corporation*, 460 F.2d 631 (8th Cir. 1972)' [1972] 1972 Washington University Law Review 804, 810.
[109] Dorothy Duffy and Marrielle B Van Rossum, 'Of Surgical Sponges and Flour Barrels, and Why Medical Experts Are Needed Even with a Res Ipsa Loquitur Instruction' [2014] 7 Drexel Law Review 309, 312.
[110] The res ipsa doctrine was established as negligence principle. At its core, the doctrine is not applicable to strict liability since the latter does not regard the defendant's conduct and care as negligence law does (even if the inferences of both areas of law effectively lead to the same result). Flowing from the technical discrepancy, several courts have established 'translations' of res ipsa-like principles into strict liability under which plaintiffs may provide circumstantial evidence, one being the malfunction theory. Owen, 'Manufacturing Defects' (n 103) 872; Beglinger (n 16), 1054.
[111] Pennsylvania is one of the many states that developed the concept of 'malfunction theory' as a circumstantial evidence principle. Some states have named it the 'indeterminate defect theory', while some others plainly refer to it as a principle of circumstantial evidence. Notably, Pennsylvania has the most developed jurisprudence on the doctrine. Owen, 'Manufacturing Defects' (n 103) 873.
[112] ibid. See also Jeffrey K Gurney, 'Sue My Car Not Me: Products Liability and Accidents Involving Autonomous Vehicles' [2013] 2013 University of Illinois Journal of Law, Technology & Policy 247, 259.
[113] *Mracek*, 610 F Supp 2d, [403].
[114] ibid., [408].
[115] ibid.
[116] ibid., [407].
[117] *Mracek*, 363 Fed Appx 925, [927] (quoting *Woloszyn v County of Lawrence*, 396 F.3d 314, 319 (3d Cir 2005))

Elmitha Pierre's proof was similarly deemed insufficient. Her circumstantial evidence to establish causation was her surgeon's differential diagnosis opinion.[118] Differential diagnosis is a process through which a physician identifies the most probable reason for a patient's condition by systematically excluding all other possible causes.[119] In essence, an expert's differential diagnosis aids the thing to speak for itself. Importantly, a differential diagnosis opinion usually centres around specific causation (as opposed to general causation), as it makes a statement on the cause of an individual patient's ailment in a particular situation.[120] Under Florida law however, an expert's injury causation analysis is insufficient if it is not, in addition to showing specific causation, supported by evidence of general causation.[121] The expert is thereby required to not only state that the product is the most probable cause for the plaintiff's injury in the case at hand (specific causation), but to also demonstrate through publicised studies, reports or scientific data that the type of injury can occur as a result of the product in question (general causation).[122]

The differential diagnosis opinion of Pierre's surgeon did not make a reference to general evidence that the da Vinci instrument could account for bowel injuries akin to Pierre's. Hence, the court deemed his testimony unreliable and legally insufficient to establish causation.[123] It did not aid Pierre's showing of general causation that she herself referred to past incidents in which the same instrument had caused arcing to patients.[124] The court stated that it would be inappropriate to consider unrelated surgeries that may have been conducted under different circumstances.[125] The district court denied summary judgement to Pierre, and properly so, according to the appellate court.[126] The latter rejected Pierre's argument anew, claiming that the surgeon's differential diagnosis could not establish causation even if viewed together with her general causation evidence that thermal injury to adjacent organs, such as the bowel could in theory arise from defective scissors.[127] The Circuit Court maintained that '[a]t best, Pierre's evidence presents a mere possibility' that an insulation defect in the Scissors caused her thermal injury.'[128] It closed the case by stating that 'a mere possibility of causation is not enough.'[129]

In both cases, the courts were reluctant to rely on circumstantial evidence. Mracek did not provide sufficient proof to effectively rule out secondary causes under the malfunction theory, and the high evidentiary standard for differential diagnosis rendered Pierre's

[118] Pierre proffered expert testimony under other claims as well all of which were deemed inadmissible for reasons of hearsay and of lack of written testimony of another surgeon now deceased. *Pierre,* 476 F Supp 3d, [1275].
[119] ibid.
[120] Thomas S Edwards and Jennie R Edwards, 'The Daubert Expert Standard: A Primer for Florida Judges and Lawyers' [2020] 94 The Florida Bar Journal 8, 13.
[121] *Pierre,* 476 F Supp 3d, [1275].
[122] Christopher RJ Pace, 'Admitting and Excluding General Causation Expert Testimony: The Eleventh Circuit Construct' [2020] 37 American Journal of Trial Advocacy 47, 47.
[123] *Pierre*, 476 F Supp 3d, [1277].
[124] ibid., [1271].
[125] ibid.
[126] *Pierre*, 2021 US App LEXIS 9623.
[127] ibid., [9-10].
[128] ibid., [11].
[129] ibid., (quoting *Hessen v Jaguar Cars, Inc*., 915 F2d 641, 647 (11th Cir 1990)).

circumstantial evidence inadmissible. In both cases, the courts deemed the evidence insufficient for the jury to construe causation.

The state of flux concerning standards of evidence under the malfunction theory and admissibility rules for differential diagnosis

In several instances, courts have held the bar high for plaintiffs to provide circumstantial evidence. Resembling Mracek's case, some courts have adopted a restrictive interpretation of the malfunction theory and required plaintiffs to effectively eliminate all other secondary reasons that could have accounted for their injuries. Such interpretation prevailed in cases surrounding complex industrial products, for example, a car's instrument panel that caught fire,[130] a bulldozer that exploded[131] or a cigarette lighter that ignited in someone's shirt pocket.[132] The plaintiffs' circumstantial evidence did not negate other potential causes to a level of sufficient probability to allow their case to proceed to the jury.

In stark contrast, other courts have shown increasing willingness to rely on circumstantial evidence.[133] Applying a broad interpretation of the malfunction theory, courts have assessed whether the harm would have ensued from an ordinary course of events. The plaintiff was not required to effectively rule out all other secondary causes, but only to show that it was more probable than not that the harm resulted from the alleged product defect.[134] Relying on this res ipsa loquitur-like idea, courts have granted summary judgement for plaintiffs injured by car accidents,[135] aeroplane crashes,[136] a burning clothes dryer,[137] or an electric blanket catching fire.[138]

[130] *Harrison v Bill Cairns Pontiac of Marlow Heights, Inc*., 77 Md App 41, 49-50 (Md Ct Spec App 1988) (affirming the grant of summary judgement to the manufacturer as 'the inference created by the doctrine of res ipsa loquitur is not applicable where the age of the vehicle, coupled with the plaintiff's failure to exclude sufficiently other causes of the accident not attributable to the defendant do not warrant the inference').

[131] *Brandon v Caterpillar Tractor Corp*., 125 A.D.2d, 627 (NY App Div 1986) (holding that a product defect could not be inferred by the sole occurrence of an accident as the defendant was not able to effectively eliminate other secondary causes).

[132] *Martin v E-Z Mart Stores, Inc*., 464 F.3d 827 (8th Cir 2006) (granting summary judgement to the manufacturer as the plaintiff had failed to evidence that the accident was not caused by wear and tear or misuse of the lighter).

[133] For further discussion on cases applying a broad application of the malfunction theory, see Bruce H Raymond and Lanell H Allen, 'Malfunction Theory as a Triple Threat for the Defense' [2020] 80 Defense Counsel Journal 297, 300.

[134] For example, the Indiana Court of Appeals stated in *Anderson v J.C. Penney Co*. that 'it is not necessary that the minds of the jurors be freed from all doubt; it is their duty to decide in [favour] of the party on whose side the weight of the evidence preponderates, and according to the reasonable probability of truth', in *Anderson v J.C. Penney Co*., 149 Ind App 325, 332-333 (Ind Ct App 1971).

[135] *Tweedy v Wright Ford Sales, Inc*., 64 Ill.2d 570 (Ill 1976) and *Moraca v Ford Motor Co.*, 666 NJ 454 (NJ 1975) (both affirming that the sole occurrence of the car accident as circumstantial evidence was sufficient to support a plaintiff's malfunction liability claim).

[136] *Lindsay v McDonnel Douglas Aircraft Corp*., 460 F.2d 631 (8th Cir 1972) (holding the manufacturer strictly liable despite that an accident report stated that the reason for the accident remained undetermined and stated three possible causes).

[137] *Cassisi v Maytag Co.,* 396 So. 2d 1140 (Fla Dist Ct App 1981) (asserting that evidence of a dryer's malfunction during normal use was sufficient proof of a defect that allowed to infer a causal connection between the machine and a house fire).

[138] *Anderson v J.C. Penney Co.,* 149 Ind App 325, 332-333 (holding that strict liability applied in a case where a child woke up in bed lit on fire by the blanket as there was reasonable probability of a malfunction drawn from the circumstances).

The body of case law on differential diagnosis displays similar ambivalence. Analogous to Pierre's case, some courts have adopted the view that a general causation opinion is a condition for differential diagnosis to be sufficiently reliable evidence.[139] Florida courts have adhered to the line of judgement and held differential diagnosis by itself to be inadmissible in cases concerning, for example, medical drugs,[140] over-the-counter dietary pills[141] or exposure to herbal pesticides.[142] Differential diagnosis would only be admissible if explicitly supported by evidence of general causation.[143]

At the same time, most other states regard differential diagnosis by itself sufficient to prove causation.[144] Some courts have even specified that showing general causation is plainly unnecessary.[145] They take the view that differential diagnosis would only be deemed unreliable if the expert's reasoning is in some way flawed or incomplete.[146] Courts have even specified that a 'lack of contextual authority' to support a differential diagnosis is not a question of admissibility, but goes to the weight of the testimony itself, which is a matter for cross-examination at trial.[147]

In sum, there appears to be little consistency – even blatant contradictions – among courts as to what constitutes sufficient and reliable circumstantial evidence for a malfunction theory to prevail or a differential diagnosis opinion to be admitted. The absence of de jure criteria in either aspect, combined with a variety of state practises erodes any predictability of the evidential requirements and outcomes of a given case. Moreover, imposing a high bar of proof is itself conceptually contradictory. On the one hand, courts have developed inference-based mechanisms to lighten the plaintiffs' burden to prove causation. On the other hand, they would refuse to rely on the proffered evidence as it encompasses too weak of a causal

[139] *Turner v Iowa Fire Equipment Company*, 229 F.3d 1202, [1205-1207] (8th Cir 2000) (holding that a physician's opinion that exposure to a fire extinguisher caused the plaintiff's airway disorder was properly excluded by the trial court as it did not rely upon specific case reports affirming such causal link).
[140] *Rider v Sandoz Pharms Corp*, 295 F.3d 1194, [1199] (11th Cir 2002) (holding that case reports containing differential diagnosis evidence did not suffice to prove a causal link between the intake of a pharmaceutical drug and the patient's symptoms).
[141] *McClain v Metabolife Int'l*, Inc, 401 F.3d 1233, [1245 – 1255] (11th Cir 2005) (finding that the trial court erred in admitting expert testimony in a toxic tort case as the expert's differential diagnosis was unreliable to establish cause without supporting evidence of general causation).
[142] *Rink v Cheminova, Inc*, 400 F3d 1286, [1295 – 1297] (11th Cir 2005) (excluding differential diagnosis as evidence for causation in a toxic tort litigation).
[143] *US Sugar Corp v Henson*, 787 So. 2d 3, [48-49] (Fla Dist Ct App 2000) (affirming the judgement in favour of the plaintiff on the ground that the differential diagnosis testimony was supported by 'generally accepted scientific principles' that exposure to the pesticides could cause neurological illnesses).
[144] The Eleventh Circuit in *Pierre* explicitly stated that 'numerous federal courts have found differential diagnosis sufficiently reliable to be admissible', as cited in *Pierre*, 476 F Supp 3d, [1277]. The federal law governing admissibility of expert evidence is Rule 702 of the Federal Rules of Evidence and the 'Daubert standard' set by the U.S. Supreme Court ruling of *Daubert v Merrell Dow Pharmaceuticals, Inc*., 509 U.S. 579, [597] (U.S. 1993). The framework set on the federal level allows states to adopt own interpretations of the admissibility and reliability on differential diagnosis evidence. Fed. R. Evid. 702.
[145] *Westberry v Gislaved Gummi AB,* 178 F3d 257, 262-63 (4th Cir 1999) (upholding a district court judgement stating that expert testimony of differential diagnosis did not require evidence of published studies or laboratory data to reliably conclude that inhalation of talc caused the plaintiff's sinus disease).
[146] *In re Paoli R.R. Yard Pcb Litig*., 35 F3d 717, 764-65 (3d Cir 1994) (holding that a differential diagnosis that failed to account for all possible causes could not be excluded except if the expert completely failed to consider other causes or provided an unjustifiably incomplete statement).
[147] *McCullock v H.B. Fuller Co*., 61 F.3d 1038, 1044 (2d Cir 1995) (noting that ambiguity surrounding the strength of differential diagnosis evidence affected the weight, not the admissibility of the testimony).

probability,[148] as in *Mracek* and in *Pierre*. Put differently, high standards of proof place barriers on the evidentiary shortcuts initially created for the very purpose of easing the plaintiff's burden. Not allowing the thing to speak for itself this way counters the res ipsa loquitur idea and repudiates policy considerations underlying strict liability.

In times of rapid technological advances, the prospective scarcity of evidence would be feasible to bridge, or at least to alleviate, with circumstantial evidence – with solutions like 'Robot ipsa loquitur.'[149] However in practice, courts manifestly tend to decide not to let the thing speak for itself if it would not spell out an almost-certainty. The lack of a coherent approach and of progressive adoption of inference-based causal reasoning may, much to the regret of future plaintiffs, turn an intended evidentiary mechanism into an increasingly elusive abstraction rooted in an archaic Latin obscurity.

Conclusion

The evidentiary filter applied in complex litigation ought to be tailored to the circumstances of the plaintiff of today and of tomorrow. This article explored how courts rejected direct evidence of defect and deemed circumstantial evidence of causation insufficient. The cases of Mracek and Pierre exemplify how, as matters become complex, some courts would consider evidence to be insufficient if it did not approximate a rather high level of probability.

Questions of causation and evidence are hardly new, rarely easy to answer, and never uncontestably. At what point the law could regard one event as to have proximately caused another has been subject to an abundance of discussion between legal scholars[150] and in courtrooms.[151] The inherent vagueness of legal causation is amplified in confoundingly complex cases in which a juror cannot rely on intuition to construe causation. The room left for ambiguity in evidential requirements appears, albeit unfortunate, a rather natural consequence. In the absence of consensus or common criteria, courts appear to impose high evidentiary standards at their own choosing. Yet, given the rapidly approaching technological developments, a robust evidentiary filter may entail problematic ramifications, two of which are notably foreseeable.

The first relates to the capacity of future plaintiffs to bring claims. If success in product liability becomes de facto conditional to proffering expert testimony, the premise is that plaintiffs have access to it. Bringing product liability claims risks translating into a question of resources – of who has the necessary connections and financial means to proffer expert evidence. In addition, future experts will need not only sufficient qualifications in their field of

[148] The Eight Circuit Court pointed to the controversy in *Lindsay v McDonnell Douglas Aircraft Corp*, 460 F2d 631, [639].

[149] Stanford University Professor Bryan Casey argues that the inference-based mechanisms for providing circumstantial evidence of causation can silence concerns about a future liability gap. See Bryan Casey, 'Robot Ipsa Loquitur' [2019] Georgetown Law Journal 62.

[150] 'Everything worth saying on the subject has been said many times, as well as a great deal more that was not worth saying' per William L Prosser, 'Proximate Cause in California' [1950] 38 California Law Review 369, 369; see also Richard Cupp, 'Proximate Cause, the Proposed Basic Principles Restatement, and Products Liability' [2002] 53 South Carolina Law Review 1085, 1086.

[151] For example, the Supreme Court of Minnesota has stated: 'There is no subject in the field of law upon which more has been written with less elucidation than that of proximate cause. Cases discussing it are legion.' *Dellwo v Pearson*, 259 Minn. 452, 453-454 (Minn. 1961).

expertise but also above-average knowledge of the technology in question. The foreseeable rise in required qualifications may limit the number of potential experts, which adds to the accessibility concerns.[152] In that way, expert evidence becoming a prerequisite may impede future product liability action and prevent plaintiffs from seeking injury compensation. Consequently, consumer protection would effectively diminish, much in contrast to the policy considerations underpinning strict liability law.

The second dimension in which high evidentiary standards prove disadvantageous is the perspective of the economic function of tort law. The purpose of fault liability law is not to point fingers for the sake of finding the culprit. Instead, the view holds that the law seeks to maximise economic efficiency by minimising costs of accidents and inducing efficient resource allocation.[153] The possibility of facing injury liability creates incentives for manufacturers to take precautions and to invest in safe innovation and continuous product development.[154] But if manufacturers are less likely to be proven answerable for risks created by their products, safe innovation may not be fostered as rigorously. It is a viable prospect that if a plaintiffs' burden remains high, the injury-deterrent function of product liability law fades.[155] The galloping progress of artificial intelligence and its expected entry in the arena of medical robotics[156] is one of the many areas that demands more, not less, impetus for safe innovation and investment.

Personal injury litigation of tomorrow may need to shift towards allowing reliance on probabilities instead of quasi-certainties – towards a weaker filter of evidence. Strict liability law ought to react to imminent societal progress to preserve its purpose as pillar of consumer protection and generator of safe investment and development. That will prevent just compensation from drifting out of reach for future plaintiffs injured by confoundingly complex machines.

[152] For further discussion on the conditions that future experts may need to meet, see Margo Goldberg, 'The Robotic Arm Went Crazy! The Problem of Establishing Liability in a Monopolized Field' [2012] 38 Rutgers Computer and Technology Law Journal 225, 247; See also F Patrick Hubbard, ''Sophisticated Robots': Balancing Liability, Regulation, and Innovation' [2015] 66 Florida Law Review 1803, 1827.

[153] See for example, Richard Posner and William Landes, 'The Positive Economic Theory of Tort Law' [1980] 15 Georgia Law Review 851, 851; Jules Coleman, Scott Hershovitz and Gabriel Mendlow, 'Theories of the Common Law of Torts', *The Stanford Encyclopedia of Philosophy* (Winter edn, 2015) <https://plato.stanford.edu/archives/win2015/entriesort-theories/> accessed 7 April 2021.

[154] ibid.; John J Donohue, 'The Law and Economics of Tort Law: The Profound Revolution' [1989] 102 Harvard Law Review 1047, 1054.

[155] An indication may be seen in the fact that after its FDA clearance twenty years ago, the da Vinci system has been linked to consistently high injury and fatality rates. Between 2004 and 2013, the number of annual injuries and deaths that occurred on average per procedure as well as the rate of broken instruments did not decrease. While these studies date back to 2015, no other studies are to indicate significant amelioration since then. Alemzadeh and others (n 2) 10-13.

[156] Cade Metz, 'The Robot Surgeon Will See You Now' *The New York Times* (30 April 2021) <https://www.nytimes.com/2021/04/30/technology/robot-surgery-surgeon.html> accessed 4 May 2021; M Mahir Ozmen, Asutay Ozmen and Çetin Kaya Koç, 'Artificial Intelligence for Next-Generation Medical Robotics' in Sam Atallah (ed), *Digital Surgery* (Springer International Publishing 2021).

Harmful but Legal: Mental Health, Liability, and the Online Safety Bill

Rachel Wilson

Introduction

In September 2021, the *Wall Street Journal* published a series of explosive articles about social media giant Facebook; the exposé revealed the extent to which the company understood how its platform inflicted harm on users, and how little they had done to address such damage. The exposé was derived from insights shared by former Facebook employee Frances Haugen, who subsequently waived her anonymity and appeared in front of a Senate committee to attest to how Facebook had, time and time again, prioritised 'astronomical profits over people.' Though egregious, the revelations came as little surprise. Numerous studies carried out in recent years have revealed the damaging effects of social media on the psychology of its users. If calls to regulate Big Tech had been gradually intensifying, they reached fever pitch after Haugen's testimony.

Regulating Big Tech is a battle that has many fronts:[1] from tax evasion to competition, attempts at limiting the power of mammoth platforms are beginning to percolate through courts in the UK, US, Europe and further afield. The European Union is actively pursuing tightened controls on Big Tech from a competition standpoint, while Haugen acknowledged that Facebook has largely appeased United States legislators for fear of greater tax penalties. Yet, the UK's Online Safety Bill is the first attempt at regulating platforms with an end to improving the user experience and preventing the damaging psychological and democratic harms that social media, in its current form, regularly imparts. First drafted following an earlier Online Harms White Paper, published in 2019, the draft Bill is currently due for amendment following evidence given to a select committee in late 2021. This 'flagship legislation' aims to bring the law up to speed where it has long been outpaced by the rate of technological innovation. Within the Bill, large providers – which will include household names such as Twitter, Instagram, Facebook, and Google – will be liable to monitor and remove damaging content from its platforms; crucially, the Bill aims at remedying not only illegal content, but legal content considered harmful to users. This includes media that is inimical to individuals, groups, and democracy as a whole. However, defining exactly *what* 'harmful but legal' means is devolved to secondary legislation.

What is remarkable about the Online Safety Bill is its wholesale acceptance that harm to users can be both physical and psychological, gradual, and incorporeal. In comparison to the status of 'psychiatric harm' at common law, which is tightly restricted to 'shocking' events,

[1] These fronts are deserving of analysis of their own and are beyond the scope of this essay, but two recent cases are worth mentioning to illustrate the scale of possible action against companies such as Facebook. In December 2021, it was reported that Rohingya victims of the Myanmar genocide were seeking to sue Facebook for £150 billion for the firm's failure to prevent the incitement of violence, while in January 2022 it was announced that the company were also facing a £2.3 billion class action in the UK on grounds of 'data exploitation' under the UK's Competition Act. Such cases are emblematic of the tipping point of sentiment calling for greater accountability and regulation.

this acknowledgement of how psychological damage is wrought constitutes a radical leap. This essay will explore how such a leap illustrates latent scepticism towards mental health in English law and whether the Online Safety Bill could give rise to civil liability, as and when it becomes an Act of Parliament.

The Provisions

The Online Safety Bill proposes imposing a three-party duty of care on search engines and 'user-to-user services' – platforms that enable users to post their own content and expose other users to such content. The services are placed in a hierarchy, with Category 1 services comprising most household names. Category 1 services are subject to the most rigorous regulation, required to protect against not only illegal content and content that is harmful to children, but a third category: content that is harmful to adults. What this term means is set out in Section 46 of the Bill, where such harms are defined as those which pose 'a material risk of . . . having, or indirectly having, a significant adverse physical or psychological impact on an adult of ordinary sensibilities.' Additionally, the Bill requires Category 1 services to prevent the removal of journalistic or 'democratically important' content. Ofcom, the regulatory and competition authority for broadcasting and telecommunications in the UK, will have the power to sanction sites that fail to protect users and review records of moderation. This last provision comes close to the 'mandated transparency' Frances Haugen has called for when attesting to the current lack of accountability social media companies face. Importantly, the detail of the definition of 'content that is harmful to adults' is deferred into secondary legislation, with the Secretary of State responsible for providing the nuance of the term and revising what content could potentially fall into Section 46's ambit.

A radical leap in understanding psychological harm

'Facebook's products harm children, stoke division and weaken our democracy.'[2] Such are the toxic effects that Frances Haugen attested to in front of a United States Senate committee in October 2021. Notably, these harms are overwhelmingly psychological; they are also *intangible* insofar as they occur online, disassociated from corporeal reality. Though they have a real-world effect, they begin as psychological germs – seeds planted whose roots take hold over time. It is impossible to pinpoint exactly when and where psychological damage truly occurs or the moment that the cognitive germ translates into physical reality and real-world action. This germination is precisely the effect platforms are designed to have; Jaron Lanier, a founder of the field of virtual reality and a former Microsoft employee, holds that the axiom: 'If you're not paying for it, you're the product' is inaccurate. In Lanier's view, 'you' are not the product: the incremental change in your behaviour over time is the product.[3] As such, any attendant detriment from behaviour modification following the consumption of content online is necessarily both temporally and corporeally vague; English law, famously, has little affinity

[2] 'Facebook products 'harm children, stoke division': Whistleblower' (*Al Jazeera*, 5 October 2021) <https://www.aljazeera.com/news/2021/10/5/facebook-products-harm-children-stoke-divisions-whistleblower> accessed 10 January 2022.

[3] Damian Whitworth, 'Social media is tearing society apart' (*The Times*, 15 November 2017) < https://www.thetimes.co.uk/article/social-media-is-tearing-societyapart-sj7km2ds7> accessed 10 January 2022.

with vagueness. In providing redress for mental inflictions, the Online Safety Bill far outreaches any precedent set at common law to date.

To illustrate how far the Online Safety Bill goes in statutory terms, we must turn to consider the chequered history of its only true counterpart at common law: psychiatric harm in relation to the tort of negligence. Psychiatric harm is narrowly, arbitrarily, and archaically defined; its limited application, compared with legal redress for physical injury, bears the influence of society's draconian attitudes and latent scepticism towards mental health and psychiatric illness. The current position in English law is that if a claimant is to recover damages for mental injury in the absence of physical injury, the plaintiff must have suffered a psychiatric injury in the form of a recognised psychiatric illness.[4] Distress alone does not qualify, and the mental injury must be intimately linked, both in time and place, to a single shocking event. In *Alcock v Chief Constable of South Yorkshire,*[5] further distinction was drawn between 'primary' and 'secondary' victims, who were directly and indirectly involved in the shocking event, respectively. Primary victims are always found to be owed a duty of care by the defendant. Secondary victims, on the other hand, are only found to be owed a duty of care if they can satisfy a high-threshold test: they must have had a 'close relationship of love and affection with the accident victim';[6] they must have witnessed the shocking event or its immediate aftermath with their unaided senses; and it was the witnessing of the event that caused a recognisable psychiatric illness. Pertinently, Lord Ackner observed that 'shock . . . involves the sudden appreciation by sight or sound of a horrifying event, which violently agitates the mind. It has yet to include psychiatric illness caused by the accumulation over a period of more gradual assaults on the nervous system.'[7] Evidently, in English law currently, no claim for psychological damages against online platforms would succeed for the very fact that any injury inflicted would be inherently disembodied.

It is instructive to examine the legal language surrounding psychiatric harm to further appreciate how significant the Online Safety Bill's acknowledgement of mental health is. Until recently, the law still employed the term 'nervous shock' – terminology introduced by medical literature in the 1870s and long since abandoned in the 1930s.[8] Here, the discrepancy between medical understandings of psychiatric health and legal understandings clarifies the gap in remedy available for victims who suffered mental rather than physical or 'bodily' harm. Damages for psychiatric harm were only available to claimants who had been involved in physical, temporal incidents, despite psychiatrists acknowledging that the 'legal prerequisite of shock, an acute emotional reaction, is an irrational limitation. It is often not a good predictor of later psychiatric illness, notably so for depression.'[9] Furthermore, the label of 'shock' betrays the law's perpetuation of the stereotype of 'instant, momentary fright as a common source of mental harm.'[10]

[4] *White v Chief Constable of South Yorkshire* [1998] 3 WLR 1509, 1518 (Lord Goff).
[5] [1992] 1 AC 310.
[6] Michael Jones, *Textbook on Torts* (8th ed., 2002) 167.
[7] *Alcock v Chief Constable of South Yorkshire,* 401 (emphasis added).
[8] Harvey Teff, *Causing Psychiatric and Emotional Harm: Reshaping the Boundaries of Legal Liability* (Hart, 2008) 8.
[9] ibid.
[10] ibid.

The Courts have historically undervalued mental injury, even while acknowledging both that the law should stay abreast of medical developments and that psychiatric damage is a confused area of the law.[11] It is now a well-established fact in medical literature that previous notions of trauma being linked solely to a single dramatic event, such as an explosion, attack, or collision, are too narrow in scope; current medical concepts of trauma embrace the idea that it can be inflicted over the course of time by exposure to difficult circumstances, such as long-term abuse, neglect, or stress. However, the law has not stayed apace, with a marked reluctance to expand qualifying tests for psychiatric harm in tort. Yet, current medical understandings of how gradual exposure to damaging circumstances can result in trauma are foundational to researchers' recognition that social media can create such psychologically damaging circumstances for its users. Despite advice to the contrary, one monolith impediment to even incrementally changing the scope of psychiatric harm is the floodgates argument. Psychiatry has been perceived as too inherently intangible, unstable and unpredictable to risk lifting the dam on claims relating to mental harm. Lord Steyn, in his judgement on *White*, at [493–494] stated: 'If claims for psychiatric harm were to be treated as generally on a par with physical injury it would have implications for the administration of justice. . . the complexity of drawing the line between acute grief and psychiatric harm [entails] greater diagnostic uncertainty [requiring] costly and time consuming [expert evidence].' Indeed, according to Sir Robin Cooke, 'nearly all arguments for restricting negligence liability are at bottom versions of the floodgates argument.'[12]

Pertinently, it was contemporary cultural understandings and awareness of mental health that informed this 'floodgates' resistance, in the way that increased literacy surrounding mental illness has now informed the Online Safety Bill. As Harvey Teff observes, at the time the Law Commission made its recommendations: 'the prospect of future mass disasters helped to generate a 'floodgates fear' which gained added momentum from extensive media coverage of PTSD as an emergent basis for claims, and from rising concern that work-related stress might become 'the next growth area."[13] Again, the language employed in the legislation and in judgments surrounding mental illness is instructive. 'Psychiatric' harm necessitates some pathological response to an event to constitute damage; the Online Safety Bill uses the non-pathological term 'psychological' and clarifies that damage can be imparted on an adult of 'ordinary sensibilities', introducing a much lower bar. It is yet to be seen whether this embrace of contemporary expertise in mental health will influence judicial attitudes towards broadening the 'psychiatric harm' definition. Nevertheless, the floodgates' fear is likely to remain.

A much wider 'floodgate'?

While the Online Safety Bill acknowledges and embraces psychological harm in a way that is unprecedented in English law, the 'floodgates' fear will nevertheless persist given the breadth

[11] *White v Chief Constable of South Yorkshire Police,* 492 (Lord Steyn stated that 'Courts of law must act on the best medical insight of the day … And psychiatric harm may be far more debilitating than physical harm.' Steyn also acknowledged the 'patchwork quilt' approach to the law, and suggested Parliament intervene to ameliorate redress. The Law Commission had published, in 1995, a recommendation of imposing a new statutory duty of care in relation to psychiatric illness, but little development occurred since).
[12] Robin Cooke, 'Tort Illusions', in P Finn (ed), *Essays on Torts* (Sydney, Law Book Co, 1989) 74.
[13] Teff (n 3) 149.

and depth of harms the Bill attempts to address. In using the terms 'duty of care' and 'adverse psychological impact', the Bill raises the spectre of possible civil action against companies: legal professionals will note that the Bill, once passed, could become fertile ground for 'breach of statutory duty' claims. However, establishing whether an Act gives rise to a statutory duty is complex and piecemeal; the law in this area is hard to clarify, and statutory legislation rarely expressly clarifies whether an imposition of duty can be made: 'the question of whether the courts will recognise such a claim is conventionally understood to be one of interpretation of the particular statute.'[14] The existence of a liability is necessarily teleological: 'whether a private right of action in tort is available to someone who is injured by the breach of a statutory duty depends upon the intention of Parliament in imposing the duty in question.'[15] If we revisit the 2019 White Paper that predated the Online Safety Bill, we find indication that the 'regulatory framework will not establish new avenues for individuals to sue companies.' Yet certain government actors have potentially vitiated this intent in subsequent comments. The current Secretary of State, Nadine Dorries, has made numerous comments to the media that the Bill will now be toughened up, with harsher sanctions for those who breach the duties of care imposed; her well-publicised remarks on criminalising trolls and social media executives who do not collaborate with Ofcom arguably blur any express intention carried in the White Paper to the contrary. Moreover, one way of parsing whether a statute might make actors liable for breach of statutory duty can be derived from the language employed; the Bill expressly imposes 'duties of care', though they may be seen to be substantively different in nature to traditional duties in tort. Given that the Bill does not expressly prevent civil action, it is more than likely litigation will be brought to test this possible tort. With the rise of 'litigation activism' and a public demand for greater accountability from Big Tech companies, test cases will surely be imminent.

Necessary teleological interpretation?

If the Bill heralds the probability of civil liability test cases, the question remains as to how the Court will interpret the Bill, once it has been fixed in statutory form. If the definition of Section 46 is not adequately sharpened, the Courts will be led into an unhelpful teleological exercise, goose-chasing Parliament's intention to determine. Since Section 46 delegates the nuanced definition of 'content that is harmful to adults' to the Secretary of State, who has the power to direct Ofcom to implement any new definitions as and when they arise, the reins that steer what conduct is acceptable or unacceptable ultimately lie in the Executive's hands. Compared to tortious 'psychiatric harm' which relies on medical literature (even if such literature is outdated), psychologically damaging content is defined with no reference to medical expertise. For this reason, the chorus of concerns that the Bill will offer a new route for state censorship is unsurprising.[16] With such powers vested to the Secretary of State, any judicial teleological interpretation could result in introducing precedents at common law that are flavoured by the

[14] James Goudkamp and Donal Nolan (eds), *Winfield and Jolowicz on Tort* (20th edn, Sweet & Maxwell, 2020) 8-001.
[15] ibid., 8-005.
[16] Alex Hern, 'Online safety bill 'a recipe for censorship', say campaigners' (*The Guardian*, 12 May 2021) <https://www.theguardian.com/media/2021/may/12/uk-to-require-social-media-to-protect-democratically-important-content> accessed 10 January 2022.

politics of the Parliament of the day; these could be enshrine a unilateral political leaning into law even though the definition of 'content that is harmful to adults' is statutorily malleable. Harms for which remedies are prioritised may correlate with the varying levels of urgencies that the UK's two political parties treat their real-world counterparts: will a government that allows rape convictions to decline so far as to essentially decriminalise the offence take rape threats made online seriously? Alternatively, will a party whose deputy leader calls the opposing party's leader 'scum' not deem right-wing journalism published on 'user-to-user services' to be 'democratically important'? It is easy to see how the ability to *define* what constitutes 'legal but harmful' in turn can constitute censorship. Current affairs will further impact the exact 'mischief' the legislation attempts to remedy. As the Bill has made its way through Parliament, such reactions have been patent: in October 2021, Nadine Dorries rushed to toughen up sanctions for Big Tech platforms and called for an end to online anonymity as a direct response to the tragic murder of David Amess MP, despite there being no proven link between online anonymity and domestic terrorism such as that fatally ended Amess' life. Such confusion over *what* can be considered damaging and knee-jerk legislation to remedy it threatens to weaken the efficacy of the much-needed legislation the Online Safety Bill seeks to introduce.

Conclusion

The Online Safety Bill is a flagship legislation; it is the first of its kind to fully attempt to treat the myriad harms inflicted by an almost entirely unregulated internet. It heralds the end of an era in which social media constitutes a 'wild west' for bullying, harassment and propaganda. Yet, given the antiquated position of our common law, it means this is the law's first foray in attempting to redress mental and emotional damage for a general population. Flaws are a given. The Bill reprises old debates around collective welfare, pitted against individual liberties: is anonymity online a right or a privilege? Must freedom of speech go unchecked if it is to truly be free? As mental wellbeing becomes a contemporary buzzword in the media, the workplace, and educational institutions alike, so too must the law respond to a need to protect people from injury – no matter how intangible and intemporal such injury is.

Bioterrorism: What Should be the Legal Response to a Global Health Threat?

Antonia Karamali

Abstract

Pandemics and bioterrorism have a lot in common, although the latter is an intentional attack whereby the former has a natural cause. In the days of Covid-19, security experts from the Council of Europe's Committee on Counterterrorism have already warned that Covid-19 may lead to the increase in the use of biological weapons by terrorists. The pandemic has shown how weak and vulnerable our modern society is to pandemics and the disruption that it could cause worldwide. Therefore, bioterrorism is one of the leading security threats of the modern world as biological attacks are indeed realistic threats to global health and should urgently trigger the interest of policy makers to adopt effective biodefense strategies. It is a real threat because biotechnology developments have made it easier, cheaper, and faster to create and transmit pathogens. This article highlights the need to strengthen the international legal framework that governs bioterrorism to improve international security. It discusses the role of science in the context of bioterrorism as well as the dual-use dilemma and the ethical questions that arise from overseeing science. Moreover, it presents the challenges of the existing international treaties and provides recommendations on the legal response to a global health threat by suggesting the establishment of a new UN legal body responsible for filling the gaps that the treaties failed to fill.

Introduction

Bioterrorism is an old practice that has existed since early civilisation and can be traced back to 430 BCE.[1] During the ancient years, biological agents (such as poisonous snakes onto enemy ships or infected blankets) used in wars were considered as unhonourable weapons and crimes against mankind.[2] However, the production of bioweapons in the modern era is synonymous with the expression of scientific and biotechnological evolution resulting in the birth of new political ambitions.[3]

There are multiple definitions of bioterrorism, either from international institutions such as the World Health Organisation or government agencies such as the Centres for Disease Control and Prevention. According to the latter, a biological attack, or bioterrorism, is the 'intentional release of viruses, bacteria, or other germs that can sicken or kill people, livestock,

[1] James W. Martin, 'History of Biological Agents as Weapons', in James R Swearengen (ed), *Biodefense Research Methodology and Animal Models* (CRC Press 2012) 2.
[2] Elizabeta Ristanovic 'Ethical Aspects of Bioterrorism and Biodefence' in Vladan Radosavljevic, Ines Banjari and Goran Belojovic (eds), *Defence Against Bioterrorism, NATO Science for Peace and Security Series A: Chemistry and Biology* (Springer 2018).
[3] Martin (n 1).

or crops.'[4] However, Nan D. Hunter, in The Law of Emergencies[5], summarised most definitions which generally include some or all of the following factors: 'the intentional use or threat of use of any biological agent to cause harm in a human, animal, plant, or other living organism; the same use or threat of use to degrade the quality of food, air, or the water supply; with the goal of influencing government conduct or policy; with the goal of intimidating or coercing a civilian population.'[6]

Unlike nuclear weapons, bioweapons cannot be easily detected as they usually do not cause recognisable symptoms of illness for several days, allowing the biological agent to spread.[7]Undoubtedly, a biological attack has the potential to cause massive disruption and seed fear and panic worldwide because societies are inextricably linked together, as millions of people cross borders every day. Therefore, establishing an international approach is vital to preparing the nations for bioterrorist attacks.

Pandemics and bioterrorism have a lot in common, although the latter is an intentional attack whereby the former has a natural cause.[8] However, bioterrorism is one of the greatest security threats of the modern world and although it should be examined from the socio-political, economic, international relations, scientific, public-health and ethical scope, this article examines bioterrorism merely from a legal perspective.

The article aims to warn that biological attacks are realistic threats to global health and should urgently trigger the interest of policy makers to adopt effective biodefence strategies.

Chapter 1 discusses the progress in biotechnology and genetic engineering and the publication of studies that may lead to the misuse of science. It highlights the importance of finding a balance between the development of biotechnology and the protection of the wider public from the dual use of science. Chapter 2 evaluates the current law governing bioterrorism and highlights the consequences of the current legal position not being amended. Finally, Chapter 3 recommends the aims and objectives of the future legal framework and provides suggestions on advancing bioterrorism prevention.

Chapter 1: Bioscience as a Security Matter

Although most nations have invested money to run scientific studies that advance biotechnology to improve modern life, the threat of bioterrorism, in which extremists use biological agents as weapons to promote political agendas, lacks equal attention. In order to understand bioterrorism, we would identify the extent to which the publication of scientific information concerning pathogens and viruses influence terrorist groups. This chapter

[4] Centers for Disease Control and Prevention: National Center for Emerging and Zoonotic Infectious Diseases, 'Anthrax as a Bioterrorism Weapon' (Centres for Disease Control and Prevention) <https://www.cdc.gov/anthrax/bioterrorism/index.html> accessed 22 October 2020.
[5] Nan D. Hunter, *The Law of Emergencies: Public Health and Disaster Management* (2nd edn, Elsevier 2018).
[6] ibid.
[7] John D. Loike, and Ruth L. Fischbach, 'Ethical Challenges in Biodefense and Bioterrorism' [2013] Journal of Bioterrorism & Biodefense <https://www.omicsonline.org/ethical-challenges-in-biodefense-and-bioterrorism-2157-2526.S12-002.php?aid=11431> accessed 22 October 2020.
[8] Alexandra Brzozowski, 'Has COVID-19 increased the threat of bioterrorism in Europe?' *Euractiv* (3 June 2020) <https://www.euractiv.com/section/defence-and-security/news/has-covid-19-increased-the-threat-of-bioterrorism-in-europe/> accessed 11 October 2020.

discusses the role of the evolving bioscience as a security issue and the need to strike a balance between scientific openness and international security in the age of bioterrorism.

The Link Between Bioscience and Bioterrorism

The effective implementation of a bioterrorist attack would require the success of several steps by the terrorist: acquiring precursor virulent biological seed cultures, growing biological agents in culture and carrying out the dissemination effectively to cause mass casualties.[9] More importantly, acquiring the dangerous microorganism to deliver a bioterrorist attack requires access and knowledge. A bio-scientist can provide both. As Vladan Radosavljevic said: 'a typical terrorist could only be a perpetrator, while a top scientist may be both, a source of agents and a potential perpetrator.'[10] Therefore, bioscience and bioterrorism are connected in this context.

Moreover, to further understand the connection between science and bioterrorism, we need to discuss the role of bioscience and the effects of 'domesticated biotechnology.' This term highlights the wide availability and assimilation of information on biotechnology to society through the publication of scientific research. According to Freeman Dyson, 'for biotechnology to become domesticated, the next step is to become user-friendly.'[11] This, in turn, poses a great threat to civilians as simplified published research on, e.g., how to create a biological agent or make a virus more transmissible may fall into the wrong hands and have catastrophic effects. As a result, there is an ethical debate over the dual use of science, research and biotechnology, a problem that has been one of the most controversial topics during bioterrorist threats.[12]

The Dual-Use Dilemma

Although the great achievements of molecular biology and genetics have advanced agriculture and medicine, they pose a potential and unpredictable risk due to the possibility that these technologies could also be used to create the bioweapons of the next generation,[13] also known as 'dual-use research of special concern.' This term refers to the research that provides knowledge, information, materials, or technologies that could be directly misused to pose a substantial threat, causing extended disaster to public health, security, agricultural plants and animals.[14] Biological weapons under the scope of the dual-use of science have become a particular area of concern because, firstly, their use can cause a serious disaster since transmission is made easier by the 'forces of globalisation' and the circulation of bodies. Secondly, the evolution of technology makes the adoption of biological agents by small groups, that lack scientific expertise and state control, easier.

[9] Charles N. Haas, 'The Role of Risk Analysis in Understanding Bioterrorism [2002] 22 Risk Analysis 671.
[10] Stephen S. Morse, 'Reflecting on Bioterrorism' [2018] 18 The Lancet: Infectious Diseases 1317.
[11] Freeman Dyson, 'Our Biotech Future' (*The New York Review*, 19 July 2007) <https://www.nybooks.com/articles/2007/07/19/our-biotech-future/> accessed 13 February 2022.
[12] Michael J. Selgelid, *Education and Ethics in the Life Sciences: Strengthening the Prohibition of Biological Weapons* (ANU Press 2010).
[13] National Research Council, *Biotechnology Research in an Age of Terrorism* (National Academies Press 2004).
[14] Sonja S Radaković, Milan Marjanović et al, 'Biological Pollutants in Indoor Air' [2014] 71 Vojnosanit Pregl, 1147.

On the other hand, the introduction of harsher laws in order to control the dual-use of research information seems inappropriate as the goal of maintaining technological supremacy, which has been an age-old argument for maximising scientific freedom and removing any limitation to scientific and technological progress, needs to be balanced over the fear that open flow of scientific research could be used by adversaries to achieve military advantage.[15] Considering the dual-use dilemma as a problem of balancing openness and profit against secrecy and security endorses the system of risk governance, which is based on efforts to forecast the future and predict the probabilities and potential costs of misuse of research.[16] As a result, the field has adopted soft law, as scientists voluntarily agreed on safety guidelines related to recombinant DNA research at the 1975 international conference in Asilomar, which became a symbol of scientific consensus and the ability for science to govern itself.[17] Consequently, practises of self-regulation, codes of conduct and risk assessment were introduced into life sciences, underlined by the language of scientific ethics and social responsibility.[18]

However, the dual-use dilemma is more complex in nature, as highlighted by the National Research Council as follows: The major vehicles of bioterrorism, at least in the near term, are likely to be based on materials and techniques that are available throughout the world and are easily acquired. The tension between the spread of technologies that protect us and the spread of technologies that threaten us is the crux of the dilemma.[19]

This spread of the technology that protects us or threatens us, can be observed in a broader scope in relation to how contemporary society has structured security. According to Foucault, 'organising circulation, eliminating its dangerous elements, making a division between good and bad circulation, and maximising the good circulation by diminishing the bad'[20]is, in fact, a vital feature of modern security governance, considering freedom and security as two complementary parts of the same system. According to Foucauldian theorists, circulation is regarded as the inherent correlate of dis-positives of security and bio-politics, and the above statement aims to highlight the distinctions between sovereign power and modern governmentality. Although the former is formulated upon the banishment of unauthorised circulation, the modern governmentality focuses on regulating circulation.[21]

Two different features can be found in the modern conception of the dual-use dilemma when considering dual-use as a matter of organising circulations. First, the positive circulation is regarded as a contribution to social and economic prosperity based on a broader neoliberal understanding of bioscience as a foundation of social and economic progress.[22] Second, it

[15] National Research Council, *Scientific Communication and National Security* (National Academy Press 1983).
[16] Dagmar Rychnovská, 'Governing Dual-Use Knowledge: From the Politics of Responsible Science to the Ethicalization of Security' [2016] 47 Security Dialogue 316.
[17] Richard Hindmarsh and Herbert Gottweis, 'Recombinant Regulation: The Asilomar Legacy 30 Years On' [2005] 14 Science as Culture 299.
[18] Albert R. Jonsen, *The Birth of Bioethics* (OUP 2003) 10.
[19] National Research Council (n 15).
[20] Michel Foucault, *Security, Territory, Population* (1st edn, Palgrave McMillan 2009) 18-19.
[21] Martina Tazzioli, 'Biopolitics of/through Mobility: The Government of Multiplicities and Processes of Value Extraction'
<https://www.academia.edu/31738317/Biopolitics_of_through_mobility_the_government_of_multiplicities_and_processes_of_value_extraction> accessed 25 October 2020.
[22] Vannevar Bush, *Science: The Endless Frontier* (US Government Printing Office 1945).

considers the use of circulation in terms of how the research is consumed. However, it is impossible to fully predict and regulate potential misuse of research and technologies because, in cases of biosecurity, not actual threats but only potential risks can be established. Since the general conception is that legitimate research is a means for progress, any legal framework that would be too broad or too restrictive seems unfit for the circumstances.[23] However, this conception poses empirical questions, such as what kind of research restrictions and what level of scientific openness would maximally promote greater good which are the next subject of discussion.

Shall We Oversee Science?

Self-regulation is an ethical dilemma for the researcher because it shifts the burden of regulation onto the scientific community and 'responsibilities', the scientists. This is because they need to prevent and foresee the potential actions and misuse of their research by others and must mitigate those risks. However, some scientists argue that overseeing research projects could not possibly refrain scientists from conducting research.[24] Rather, they argue that not regulating the field poses a greater risk to science since a lab accident is enough for policymakers to impose draconian restrictions because of the accident. Consequently, that reaction will not be focused on the very small group that is responsible for the accident but on bioscience.[25]

On the other hand, it is possible that researchers may not always willingly put security interests above their own career advancement. There is the argument that scientists should not play the role of the police as they have not been trained to do so and that the scientific community should not be held liable for effects that are beyond their control.[26] Therefore, the scientific community and policymakers need to cooperate in a new way, given the fact that censorship goes against autonomy which is essential to scientific advancement.

Although it is hard to provide a definite answer to the question on scientific oversight, the focus should be on the introduction of sound and balanced policies that would not undermine international welfare, contribute to systemic inequities and injustices, and violate the trust that the public places in the science community to best serve its collective interests[27]

The Way Forward

Since this article aims to answer the question of what the legal response to a global health threat occurred by a bioterrorist attack should be, the following recommendation provides a solution regarding the transformation of the governance system in order to meet the new security challenges in terms of the dual-use dilemma which contributes to the threat.

[23] Hindmarsh and Gottweis (n 17).
[24]Wendee Holtcamp, 'One Study, Two Paths: The Challenge of Dual-Use Research' [2012] 120 Environmental Health Perspectives A241.
[25] ibid.
[26] Hanz-Jörg Ehni, 'Dual Use and the Ethical Responsibility of Scientists' [2008] 56 Arch Immunol Ther Exp (Warz), 147.
[27] Joseph R. Biden Jr, 'Memorandum on Restoring Trust in Government Through Scientific Integrityand Evidence-Based Policymaking' (*The White House: Briefing Room* 27 January 2021) < https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/27/memorandum-on-restoring-trust-in-government-through-scientific-integrity-and-evidence-based-policymaking/> accessed 9 January 2022.

The need to oversee knowledge is a new task that needs to be fulfilled and is linked to the development of knowledge societies, which are highlighted by the monetisation of knowledge and fading exclusivity of scientific expertise. A combination of self-governance upholding ethical norms following responsible science and sovereign power may achieve the balance between open science and security. Therefore, this article recommends the establishment of an international authority within the World Health Organisation (WHO) responsible for overseeing research projects to promote confidence that research is scientifically justified, conducted in a safe and secure manner. Currently, WHO only oversees work on the smallpox virus, but it is recommended to extend this diacritical oversight to all research projects internationally.[28]

Furthermore, continuous education and training provided by the WHO about the responsibilities of scientists regarding dual-use research are the key tools to govern scientific knowledge effectively. Finally, the strengthening of The Biological and Toxin Weapons Convention to prevent misuse by controlling the access to dual-use technologies and materials or giving the WHO authority to require licensing of those using such technologies and materials, might be the way forward.

Chapter 2: International Treaties Governing Bioterrorism

Having addressed the role of science in the age of bioterrorism, the next matter that needs to be addressed is whether the law on biological weapons is adequate to prevent bioterrorism. The most important legal piece is the Biological and Toxin Weapons Convention (BWC), which prohibits the use of biological agents and toxins that cause harm to humans, animals, or plants. This chapter highlights the challenges that the Convention needs to address to fulfil its aims. Moreover, it discusses the role of UNSCR 1540 and its contribution to filling the gaps of BWC.

The Ambiguity of BWC

The BWC is a cornerstone of the framework prohibiting weapons of mass destruction which came into force in 1975 and was the first multilateral disarmament treaty prohibiting the development, production, acquisition, transfer, stockpiling and use of biological and toxin weapons.[29]The BWC currently counts 183 States Parties and four Signatory States. However, there are ten States which have neither signed nor acceded to the Convention.[30]

BWC is a succinct agreement, and the brevity illustrates this in the convention provisions. However, as a result of its brevity, the provisions are ambiguous, and consequently, the obligations of States parties under the convention on matters concerning the States' compliance with the BWC and their expectations of each other, have not been made clear.

Furthermore, the BWC's scope of prohibited activities also lacks clarity. It is difficult to determine what is prohibited from being used because Article I only provide a broad definition of biological agents, as the treaty does not include a list of specifically prohibited

[28] Ron Fouchier and AB Osterhaus, 'The Fight Over Flu' [2012] Nature 257.
[29]Graham S. Pearson, 'Time for Structural Changes to Make the Biological and Toxin Weapons Convention More Effective' [2016] 1 Global Security: Health, Science and Policy 23.
[30] United Nations: Office for Disarmament Affairs, 'Achieving Universality' (United Nations) <https://www.un.org/disarmament/biological-weapons/about/universalization-and-joining-the-bwc/> accessed 27 October 2020.

agents or quantities. This is also relevant to the dual-use problem mentioned in Chapter 1, as the production of biological agents has the potential to be used as weapons but have peaceful purposes as well.

BWC does not pose a blind prohibition on all biological weapons, but it bans 'types' and 'quantities' of biological agents that have 'no justification for prophylactic, protective, or other peaceful purposes.'[31] Although this broad provision may create a 'future proof' treaty that has the mechanism to follow scientific development, the lack of definition on what constitutes 'peaceful purposes' creates a legal regime with no specificity. Attempts to reach an agreement on an approved list with types of prohibited materials have been made in the past but have failed. Therefore, the state's parties are free to decide themselves what types or quantities have no reasonable purpose, leaving open space for manipulation of the standards.[32]

Moreover, Article V of the BWC imposes an obligation on all states parties to take measures to adjust their national laws in order to comply with the Convention. However, due to the lack of guidance as to how the Convention should be implemented, only a small number of states parties have enacted national legislation in accordance with this provision of the BWC.[33] As a result, the BWC raises concerns over its compliance since history has shown that many states, such as the former Soviet Union and Iraq, have developed after the BWC enactment, major clandestine bioweapons programmes[34] as well as many countries continuing to produce biological agents claiming 'research purposes.' It seems that many state parties have taken lightly their obligations imposed by the BWC and therefore, more work needs to be done to enhance its compliance mechanisms and implementation.

The BWC has been criticised for failing to adopt a verification and inspection protocol, which was being negotiated to improve the effectiveness of the Convention.[35] Therefore, States focused on developing preventative codes as a priority rather than developing a compliance protocol. However, Article V requires 'The States Parties to this Convention to consult one another and to cooperate in solving any problems which may arise in relation to the objective of, or in the application of the provisions of, the Convention.' Consequently, it can be said that the 'compliance' or 'enforcement' mechanisms included in the text of BWC have only a 'consultative' nature.

In case of a complaint, Article V allows for a 'formal consultative meeting' (FCM) in order to consider an allegation of noncompliance.[36] This would lead to an investigation by the Security Council who will obtain information by the Parties as no independent report on the complaint is authorised. The Security Council has the authority, under Chapter VII of the UN

[31] United Nations: Office for Disarmament Affairs, 'Science and Technology under the Biological Weapons Convention' (United Nations) <https://www.un.org/disarmament/biological-weapons/science-and-technology/> accessed 3 November 2020.
[32] Daniel H. Joyner, *International Law and the Proliferation of Weapons of Mass Destruction* (OUP 2009) 91-92.
[33] Masahiko Asada, 'Security Council Resolution 1540 To Combat WMD Terrorism: Effectiveness and Legitimacy in International Legislation' [2008] 13 Journal of Conflict & Security Law 303, 306-07.
[34] John R. Walker, 'Strengthening the BTWC' [2003] 4 EMBO Reports: Science & Society S61.
[35] Marie Chevrier, 'The Biological Weapons Convention: The Protocol that Almost Was' in Trevor Findlay and Oliver Meier (eds), *Verification Yearbook* (Verification Research, Training and Information Centre, 2001).
[36] Angela Woodward, 'The BWC and UNSCR 1540' in Peter van Ham & Olivia Bosch (eds), *Global Non-Proliferation and Counter-Terrorism* (Brookings Institution 2007), 103-04.

Charter, to act against a state party that violates provisions of the BWC. However, this is unlikely as it is very easy for the party in question to hide evidence or restructure its practises so that they meet the requirement of peaceful purposes, making it impossible for a complaint to result in an action being taken against the state.

The lack of sufficient verification mechanisms led to the adoption of various voluntary confidence-building measures (CBMs) that called for the exchange of information about research centres and laboratories with high-containment facilities and data on unusual outbreaks of disease.[37]Under the current CBM regime, states parties must report to all other states; data on laboratories and research centres, national biological defence research and development as well as past activities in offensive or defensive biological research. Moreover, they must report efforts to encourage publication of results of biological research directly related to the BWC and declaration of legislation or other measures states have taken to implement the BWC.[38] However, once again most states do not follow their obligations to exchange information.

The revolution in biotechnology does not allow the states parties to take lightly their obligations imposed by the treaty. Therefore, a re-drafting or re-envisioning of the disarmament imperative embodied in the BWC and clarification on the ambiguities mentioned above must be addressed urgently. If nothing is done, the indeterminate provisions and a lack of transparency pose serious risks for future undetected defections. Furthermore, history shows that language that fails to regulate related attractive weapons systems had the effect of channelling arms races into those unregulated areas rather than ending the arms races themselves.[39] Therefore, if the BWC is not amended, the Convention may have an opposite effect to the ones expected.

In addition, a package of verification measures such as the requirement of declarations or on-site inspections, sampling and identification need to be agreed upon to strengthen the BWC. Since no actions for violation have been made to the Security Council so far, states parties need to start reporting cases of non-compliance to the Security Council so that the effectiveness of the 'standard' enforcement regime established in BWC can be tested and reformed as appropriate.[40] Moreover, greater 'universalisation' of the BWC is required so the Middle Eastern States, such as Israel, Egypt and Syria that haven't joined the BWC, sign the treaty.

In conclusion, the fact that some states consider biological weapons as a low-security priority but regard biotechnology as significant in responding to social challenges has resulted in an increasingly incrementalistic process in the development of the BWC with veto powers and diplomatic procedures.[41] The treaty seems to be 'muddling through' rather than following

[37] Jack M. Beard, 'The Shortcomings of Indeterminacy in Arms Control Regimes: The Case of the Biological Weapons Convention' [2001] 101 American Journal of International Law 271, 282.
[38] 'Third Review Conference of the Parties to the Convention on the Prohibition of Development, Production and Stockpiling of Bacteriological (Biological) and Toxin Weapons and on their Destruction' (Geneva, 9-27 September 1991) UN Doc BWC/CONF.III/23.
[39] Beard (n 37), 290.
[40] Michael Moodie, 'The Soviet Union, Russia, and the Biological and Toxin Weapons Convention' [2001] The Non-Proliferation Review 59, 59-61.
[41]James Revill, 'Muddling through' in the Biological & Toxin Weapons Convention' [2018] 55 International Politics 386, 399.

any form of paradigm shift.[42] In fact, those optimistic states parties wishing to enhance the convention find incrementalism the only approach to building the biological weapons regime, especially in this case, where power is institutional, and the interests of states vary considerably.[43]

The Role of UNSCR 1540

The uniqueness of the bioterrorist threat made the states negotiate a legally binding instrument, under the UN mandate, to check Non-State Actors (NSAs) and limit their role in bioweapons and terrorism. As a result, the United Nations Security Council Resolution 1540 (UNSCR 1540) was established in 2004. Although BWC covered the actions of NSAs, The UNSCR 1540 filled the gap in the international non-proliferation regime, which lacked the legal basis to prosecute NSAs that were engaged in the proliferation of Weapons of Mass Destruction.[44] The Resolution called on all UN members to introduce legally binding measures for criminalising proliferation.[45]

Although it seems to be a measure against proliferation, the Resolution significantly prevents bioterrorism. In the bioterrorism context, UNSCR 1540 goes beyond the BWC because it focuses on non-state actors. Having the effect of a UN Security Council Resolution and binding states, not parties, to the BWC provides specificity regarding measures states must take to help prevent bioterrorism and makes a step in the direction of a quasi-compliance body with limited verification and enforcement role.[46]

However, the dynamic element of the Security Council Resolution that binds all nations has generated great criticism and is seen as a possible weakness. Some scholars, such as Professor Masahiko Asada, have challenged the legality of this international legislation because the legal rules stated in UNSCR 1540 have been drafted by a limited number of states.[47] In fact, legislation by the Security Council means 15 states establish general rules that legally bind 192 members of the United Nations, meaning that the international community would need to follow the rules without participating in their drafting process or having the opportunity to debate on their appropriateness. Therefore, some states have argued that the Security Council lacks the competence to consider all parties' interests, and some states may refuse compliance.

The above arguments raise questions on freedom and sovereignty. Usually, the states have the freedom to opt-in or refuse participation in a treaty. As a result, this freedom allows the states to protect and ensure their national interest and sovereign rights. However, Security Council legislation does not allow such sovereign freedom.[48] Moreover, as discussed in the

[42] ibid.
[43] ibid.
[44] Huma Rehman and Afsah Qazi, 'Significance of UNSCR 1540 and Emerging Challenges to its Effectiveness' [2019] 39 Strategic Studies 48, 50.
[45] ibid.
[46] UNSC Res 1540 (28 April 2004) UN Doc S/RES/1540.
[47] Asada (n 33), 322-323.
[48] ibid, 324-325.

case of *The Prosecutor v Tadić*,[49] the Security Council has broad discretion in deciding on the course of action and evaluating the appropriateness of the measures to be taken in case of violations.[50] Although this means that the states will be held accountable for their actions, a more codified legal regime regarding the Council's discretion will provide more certainty and trust between states and the Council.

Although UNSCR 1540 does not establish an independent international body for compliance and does not provide a list with what specific protection measures must be taken to meet international standards, it establishes a system under which states have obligations in various areas concerning biosecurity and law enforcement. Consequently, states are required to legislate in all areas concerning biosecurity and have an obligation to enforce such legislation, creating domestic norms.

Other International Legislation

Although there are no other major pieces of legislation affecting bioterrorism, alternative instruments that work as confirmatory to the rules established by the mechanism mentioned above. For example, the Terrorist Bombing Convention criminalises the unlawful use of 'explosives and other lethal devices, including' toxic chemicals, biological agents or radiation or radioactive material.[51] Furthermore, transportation-related treaties also aim to control bioterrorism by criminalising the transportation of various biological weapons, in civil aviation through the Beijing Convention[52] and maritime SUA Protocol modes.

However, no treaty enforced so far has managed to fill the gaps in vital matters such as the lack of definition of what constitutes a bioweapon or which agents and what quantities are banned, or the establishment of adequate verification, inspection and enforcement mechanism. Treaties governing bioterrorism need to mature throughout the years by learning from past experiences and revising their approaches accordingly to ensure their continued contemporary relevance and effectiveness.[53] However, all deficiencies mentioned above, as well as the non-uniformity of compliance, are long-term challenges that the international community must address by providing specific solutions to specific problems.

The 'Balkanisation' of Law on Bioterrorism

The absence of an effective unified international mechanism for coordination in bioterrorism increases instability. This leads to the balkanisation of the legislation that governs bioterrorism. It has been argued that the phenomenon of 'balkanisation' appears to be relevant at a time when 'global institutions are weakening, and the tools for collective international action are fracturing.'[54] Balkanisation is a modern metaphor used to describe territorial fragmentation of

[49] *The Prosecutor v Tadić* (Decision on the Defence Motion for Interlocutory Appeal on Jurisdiction) [1996] 35 ILM 32.
[50] ibid., [28].
[51] Terrorist Bombing Convention (adopted 15 December 1997 UNGA Res 52/164) art 1[3b].
[52] Michael Jennison, 'The Beijing Treaties of 2010: Building A 'Modern Great Wall' Against Aviation-Related Terrorism' [2011] 23 Air & Space Law 9, 10.
[53] Rehman and Qazi (n 44), 66.
[54] James Andrew Lewis, 'Sovereignty and the Evolution of Internet Ideology' (*Center for Strategic and International Studies* 2020) <https://www.csis.org/analysis/sovereignty-and-evolution-internet-ideology> accessed 25 November 2020.

a region into smaller non-cooperative states.[55] However, it has been widely used in data policy arguments and has been extended to other fields such as law and security studies.[56] Many officials have argued that bioterrorism needs a collective approach and, although the negative outcome of balkanisation in this area is significant, implementation of a common set of rules is far from a reality due to the existing treaties not providing such a set of rules.

Moreover, the states have many economic, socio-political, and constitutional differences and it should be acknowledged that there can be no single legal system that would suit and perfectly serve all countries. As a result, most states that do not wish to compromise have extended their sovereign control and have introduced their own rules for preventing and countering bioterrorism whereby the protection of their national interests and citizens has been prioritised, abandoning the call for a unified international approach. Consequently, due to the increased shift towards national regulation of bioterrorism, we have witnessed the departure from a 'unified international system' to the balkanisation of the law pose barriers which restricts the treaties from reaching their full potential.

For all the reasons mentioned above, and due to core issues surrounding the implementation and enforcement of treaties, most of them have a consultative role and are treated as 'soft law.' Although the soft law approach may seem beneficial in other contexts (as seen in Chapter 1 in terms of regulating science), the long-term impact of its rational mechanism in controlling bioweapons and multilateral security instruments presents fundamental problems that pose different research agendas.[57]

Chapter 3: The Legal Response to Bioterrorism

Having discussed the core international treaties that govern biological weapons and their deficiencies, it is evident that existing prevention strategies and laws are insufficient to guarantee that biological weapons will not be used against civilian populations.[58] This conclusion is drawn from the fact that there are still countries such as Iran and Syria that seek to possess or produce biological weapons[59] and there are countries like Israel and UAE that have not yet signed the BWC.

A thoughtful analysis of the consequences of a bioterrorist attack provides a mandate for action.[60] Therefore, the international community must take collective action to prevent such attacks from happening. At this stage, there is no international organisation, in the United Nations system or elsewhere, responsible for overseeing biosafety and biosecurity at a global

[55] Nikolina Bobic, *Balkanization and Global Politics: Remaking Cities and Architecture* (1st edn, Routledge 2020).
[56] ibid.
[57] Asada (n 47), 321.
[58] Thomas V. Inglesby, Tara O'Toole and Donald A. Henderson, 'Preventing the Use of Biological Weapons: Improving Response Should Prevention Fail' [2000] 30 Clinical Infectious Diseases 926, 927.
[59] General Intelligence and Security Service (AIVD), 'Proliferation of Weapons of Mass Destruction: Risks for Companies and Scientific Institutions' (*Federation of American Scientists*, July 2003) <https://fas.org/irp/world/netherlands/wmdrisks.pdf> accessed 28 December 2020.
[60] Philip K. Russell, 'Biologic Terrorism: Responding to the Threat' [1997] 3 Emerging Infectious Diseases 203, 204.

level.[61] Since one of the article's aims is to fill that gap, Chapter 3 will, on one hand, demonstrate the need for a much higher level of global governance of biosecurity. Then, on the other hand, it will provide recommendations for the establishment of a new international legal framework on bioterrorism within the United Nations, referred to in this Chapter as the 'UN legal body.'

Bioterrorism Requires a Unique Approach

Chemical, biological, and nuclear weapons have different scales of harm and rely on different physical principles but 'weapons of mass destruction' puts them all together.[62] Consequently, policymakers incorrectly equate and generalise the elements of weapons of mass destruction and that leads to inadequate countermeasures.

Biological weapons present a fundamentally different challenge from nuclear and chemical weapons because in the case of a bioterrorist attack, first responders would be the public health and healthcare systems, not firefighters, law enforcement, and emergency response personnel.[63] The reason for this is because a covert bioterrorist attack would likely come to our attention gradually, as doctors became aware of an accumulation of unexpected illness or death among previously healthy people. The speed and accuracy with which physicians and laboratories reach correct diagnoses and report their findings to public health authorities would directly affect the number of deaths and - if the attack employed a contagious disease - the ability to contain the epidemic.[64]

Therefore, scholars suggest that the international community should recognise the different scale in the degree of threat posed by various forms of 'weapons of mass destruction.'[65] A strong strategy must relate the costs and dangers of countermeasures to the scales and probability of the threat being countered.[66] If the distinction between the different forms of weapons of mass destruction is not properly appreciated, the security risks of action taken may outweigh the risks of the targeted threat[67] resulting in weak preparedness.

Moreover, bioterrorism requires a different approach from other weapons of mass destruction because it poses deep legal challenges, as it needs to bring diverse areas of law into play to tackle a public health emergency. For example, bioterrorism combines criminal law as the use, development, or possession of a biological weapon by a state is already a crime under

[61] Jing-Bao Nie, 'In the Shadow of Biological Warfare: Conspiracy Theories on the Origins of COVID-19 and Enhancing Global Governance of Biosafety as a Matter of Urgency' [2020] 17 Journal of Bioethical Inquiry 1.
[62] Zachary Kallenborn, 'In Defence of 'WMD': A War of Words and the Challenge of Swarms (*War on the Rocks*, 23 July 2020) <https://warontherocks.com/2020/07/in-defense-of-wmd-a-war-of-words-and-the-challenge-of-swarms/> accessed 19 January 2021.
[63] David P. Fidler, 'Bioterrorism, Public Health, and International Law' [2002] Chicago Journal of International Law Paper 7.
[64] Tara O'Toole, 'Hearing on Terrorism Preparedness: Medical First Response' (Johns Hopkins: Center for Health Security, 22 September 1999) <https://www.centerforhealthsecurity.org/our-work/publications/hearing-on-terrorism-preparedness-medical-first-response> accessed 15 January 2021.
[65] ibid.
[66] Joseph Cirincione, Jessica Tuchman Matthews et al, 'WMD in Iraq: Evidence and Implications' (Carnegie Endowment for International Peace 2004) <https://carnegieendowment.org/2004/01/08/wmd-in-iraq-evidence-and-implications-pub-1435> accessed 15 January 2021.
[67] ibid.

international law subject to the principle of universal jurisdiction.[68] It also combines National Security Law since a public health crisis is a threat to the nation; Public Health Law as preparedness for bioterrorist attacks requires the existence of a legal framework for effective coordinated responses to emergencies; Human Rights Law because measures to respond to a biological attack may result in infringement of individual rights and liberties such as the freedom of movement. Finally, it combines International Law, amongst others, because international regulations are needed in order to ensure accurate data and early reporting of a global infectious disease.

In conclusion, bioterrorism poses different risks from other weapons of mass destruction and therefore, the biodefense approach should be proportionate to those risks. Moreover, due to the complex nature of bioterrorism, further developments to boost preparedness and effective response should include the cooperation of all relevant instruments that contribute to implementing a strong biodefense strategy such as the science community, epidemiologists, policymakers and law enforcement officers, amongst others.

The Need to Modernise the International Legal Framework

This article does not suggest that there is a perfect system to safeguard against terrorist use of a biological agent. Rather, it argues that conscientious preparation — to the greatest extent that budgets and available scientific methods allow — will reduce the probabilities or consequences of an actual attack.[69] In order to do that, there is a great need to modernise the current international laws and policies that govern bioterrorism.

In order to understand the need to modernise the law on bioterrorism, it is necessary to analyse the socio-political characteristics that shape our society. A key conclusion from sociological theories of risk is that modern societies not only create new risks by advancing technologies, but also want to control them.[70]

According to Beck, the focus of modern global risk societies is on preventing any bad (though often unmanageable) events.[71] He argues that 'reflexive modernity' is accompanied by the heightened sense of uncertainty and vulnerability[72] as an enduring feature of modern culture.[73] Furthermore, both Beck and Giddens argue that increased knowledge leads to progress, but this knowledge has created new hazards that we do not possess the ability to understand and manage effectively by enforcing the existing laws.[74]

Moreover, when arguing about modern forms of risk, Aradau and Van Munster also state that 'the rationality of catastrophic risk translates into policies that actively seek to prevent

[68] Michael P. Scharf, 'Clear and Present Danger: Enforcing the International Ban on Biological and Chemical Weapons Through Sanctions, Use of Force, and Criminalization' [1999] 20 Mich J Intl L 477, 506.
[69] National Research Council (US) Panel on Biological Issues, *Countering Bioterrorism: The Role of Science and Technology* (National Academies Press 2002).
[70] Colin McInnes and Anne Roemer-Mahler, 'From Security to Risk: Reframing Global Health Threats' [2017] 93 International Affairs 1313.
[71] Ulrich Beck, *Risk Society: Towards a New Modernity* (Sage Publications 1992), 10–21.
[72] Ulrich Beck, Wolfgang Bonss and Christoph Lau, 'The theory of reflexive modernization: problematic, hypotheses and research programme' [2003] 20 Theory, Culture and Society 1.
[73] ibid.
[74] National Research Council (n 15); O'Toole (n 84).

situations from becoming catastrophic at some indefinite point in the future.'[75] This applies to bioterrorism since attacks with biological agents are expected to cause major impact and severe damage - prevention must be employed to avert their disastrous effects.[76] Therefore, bioterrorism and public health should be examined by analysing the socio-political context of contemporary societies since the response to a bioterrorist attack is incorporated into their structures and is a direct consequence of the developments, changes, and orientations of governments and modern societies.[77]

Another reason that justifies the need to modernise bioterrorism law is that it is antiquated. Most public health laws have not been systematically updated since the early to mid-20th century, with very few treaties added in the last two decades. As a result, most nations have developed their own laws that are inconsistent in their structure, substance, and procedures for detecting, controlling, and preventing bioterrorist attacks, resulting in legal inconsistencies concerning the same issues.[78] As discussed in Chapter 2, this goes against the unified approach in tackling bioterrorism that the Treaties aimed to achieve. Finally, the law should follow the evolution of technology as new biotechnology poses new risks.[79] Particularly, the development of the synthetic biology could give terrorist groups the ability to synthesise dangerous pathogens more easily than in the past.

Recommended International Legal Framework

The approach against bioterrorism is a global concern and necessitates that the issue should be criminalised and dealt with internationally, with the assistance of international cooperation.[80]Therefore, this article suggests the establishment of a UN legal body.

The first step to determine the legal framework is to decide the function of the new legal framework. For example, deterring groups and states from developing and using bioweapons or preparing and empowering governments and private institutions for responding to a biological attack could be positive developments. Furthermore, another function could be the regulation of exercise of government power in public health emergencies to ensure the protection of individual rights and freedoms. Finally, facilitation of attribution of biological weapon attacks and retribution against the perpetrators are all functions that the new legal strategy should seek to implement.[81]

However, what would make the new legal UN legal body unique is that all these functions would be common to all nations collectively at a global level and the new body would fill the gap of not having an international institution charged with overseeing biosecurity

[75]Claudia Aradau and Rens van Munster, 'Governing Terrorism Through Risk: Taking Precautions, (un)Knowing the Future' [2007] 13 European Journal of International Relations 89, 105.
[76] Radaković (n 14).
[77] ibid.
[78] Lawrence O. Gostin, and James G. Hodge Jr., 'Public Health Emergencies and Legal Reform: Implications for Public Health Policy and Practice' [2003] 18 Public Health Reports 477.
[79] Interpol, 'International Experts Meet on Potential Threat Posed by New Technologies (*Interpol*, 5 December 2018) <https://www.interpol.int/News-and-Events/News/2018/International-experts-meet-on-potential-threat-posed-by-new-technologies> 19 January 2021.
[80] Divashree Sharma and Ambrish Mishra et al, 'Bioterrorism: Law Enforcement, Public Health & Role of Oral and Maxillofacial Surgeon in Emergency Preparedness' [2016] 15 Journal of Maxillofacial and Oral Surgery 137.
[81] Hunter (n 5).

globally. Therefore, an international organ with legal powers built within the United Nations seems the most effective response to fill that gap.

The suggested objectives of this new legal system would be to interpret the relevant laws to evaluate whether they provide the needed authority for responding to the threat posed by bioweapons.[82] It would modify existing law or create new ones to strengthen the legal regime supporting biodefense.[83] Moreover, it would coordinate legal authorities among governmental departments of the nations[84] and implement the legal authority properly through adequate funding, personnel, training, technology, and equipment,[85] covering most of the known gaps.

The Recommended Model

Since the functions and objectives are set, the next step is to determine the model that this new UN body would adopt. While there is widespread agreement that the public health system will play a crucial role in the event of a bioterrorist attack,[86] there is much debate as to how the system should respond to such a catastrophic event.[87] Multi-level governance in global health poses many complexities thus an appropriate system that would allow the effective function of the body needs to be developed.

Federalism is a type of political system in which the advantages of shared rule are combined with those of regional government.[88] Federal systems allow flexibility and provide clear distinctions of the regions within a nation and allow for region-specific policy approaches to be developed.[89] However, federal systems pose limitations because the division of powers can create an obstacle to the development of centralised approaches to international challenges and that would lead to a slow response in a bioterrorist emergency. If the recommended UN body were to adopt a federal system, its limitations would be confused when international treaties are signed by the parties of the UN as its implementation and compliance would require the cooperation of national and regional governments.[90]

Moreover, a system based on the Rule of Law would also seem problematic in the case of biological attacks. According to the Rule of Law, human affairs are governed by law, not the arbitrary exercise of power.[91] However, in the case of a health emergency, states find themselves under immense pressure, resulting in the abandonment of fundamental rights

[82] RE Hoffman and JE Norton, 'Lessons Learned from a Full-Scale Bioterrorism Exercise' [2000] 6 Emerg Infect Dis 2000 652.
[83] National Commission on Terrorism, 'Countering the Changing Threat of International Terrorism' (*FAS Intelligence Resource Program*, June 2000).
[84] O'Toole (n 84).
[85] Hunter (n 5).
[86] Commission on the Prevention of WMD Proliferation & Terrorism, 'World at Risk' (*Library of Congress*, 2008) <www.preventwmd.gov/static/docs/report/worldatrisk_full.pdf.> accessed 12 January 2021.
[87] Sam Berger, 'Public Trust, Public Health, and Public Safety: A Progressive Response to Bioterrorism' [2010] 4 Harvard Law & Policy Review 295, 298.
[88] Ronald L. Watts, *Comparing Federal Systems in the 1990s* (Queen's University 1996) 6–14.
[89] Kumanan Wilson, Christopher McDougall and Ross Upshur, 'The New International Health Regulations and the Federalism Dilemma' [2005] 3 PLoS Med e1.
[90] ibid.
[91] Donald A. Henderson, Thomas V. Inglesby, Tara O'Toole and David P. Fidler, 'The Malevolent Use of Microbes and the Rule of Law: Legal Challenges Presented by Bioterrorism' [2001] 33 Clinical Infectious Diseases 686.

established by the Rule of Law.[92] The insufficiency of current legal frameworks in case of bioterrorist attack would put pressure on nations to take draconian measures to tackle the crisis that might undermine the Rule of law, leading to social unrest and distrust. In addition, measures to contain an epidemic usually require violation of individual rights through such acts as forced quarantine, isolation etc. Therefore, the Rule of Law could create a challenge to the recommended UN legal body to find the balance between public health actions and civil rights. The biggest challenge would be to use the Rule of Law as a weapon rather than a weakness in tackling biological attacks.

The Rule of Law can be used as a weapon because it promotes transparency, participation, representation, equality, and accountability, among other principles.[93] These elements look great in theory but, in a case of a health emergency, engagement and representation are not feasible. Therefore, the biggest challenge with the Rule of Law would be legality. In health emergencies, when Parliaments struggle to cope because of the measures in place, laws are passed fast and without much debate. Consequently, there have been cases where there was a lack of coherence to clarify which body was competent to enact measures and on which legal basis.[94] Therefore, following the Rule of Law by the book in case of health emergencies would result in a slow process when rules are needed urgently.

From my point of view, neither federalism nor the Rule of Law is the answer to an effective legal model to counter bioterrorism. The existing legal systems do not cover the challenges that bioterrorism poses. Therefore, to rectify the deficiencies in current law and find a balance between individual rights and effective public health actions a coordinated international corporation and legal action is needed.[95] This is a Herculean task that will develop a new legal model which may contain features of the existing systems but combined with fresh ideas that will evolve the way we respond to bioterrorism. An attempt to provide recommendations on how we could advance the way we respond to bioterrorism is made below.

The New UN Legal Body in Practice

Even the most sceptical would agree that bioterrorist attacks are unpredictable, their consequences unavoidable and potentially catastrophic, and the answer to an effective policy response is strengthening preparedness. Preparedness for bioterrorism will also strengthen the response to naturally occurring epidemics.[96]Moreover, effective preparedness is, in itself, a deterrent to bioterrorism since it reduces the incentive to use biological weapons by making a country or region a hard target.[97]

[92]ibid.
[93] Dr Julinda Beqiraj, Dr Jean-Pierre Gauci and Nyasha Weinberg, 'The Rule of Law in Times of Health Crises' [2020] British Institute of International Comparative Law <https://www.ed.ac.uk/global-health/resources/papers-presentations/rule-of-law-in-times-of-health-crisis> accessed 19 January 2021.
[94] DRI, 'The Rule of Law Stress Test – EU Member States' Responses to Covid-19' (*Democracy Reporting International*, 19 May 2020) <https://democracy-reporting.org/en/office/EU/publications/the-rule-of-law-stress-test-eu-member-states-responses-to-covid-19> accessed 10 January 2021.
[95]ibid.
[96] Manfred S Green, James LeDuc, Daniel Cohen and David R Franz, 'Confronting the Threat of Bioterrorism: Realities, Challenges, and Defensive Strategies' [2019] 19 The Lancet Infectious Diseases e2.
[97] ibid.

Therefore, the recommended UN body could seek to fill the gaps that the BWC and UNSCR 1540, as discussed in Chapter 2, could not fill. These gaps include the verification issue above, the issue of a meaningful definition of biological weapons and the lack of enforcement mechanism for violation. They also include the lack of funding for training, a lack of development and coordination of bioterrorism forensics capabilities and lack of business-level regulations. Moreover, the establishment of 'in-house' intelligent services for bioterrorism within the UN body, the establishment of a special committee that would investigate the potential use of biological weapons are some problems that remain unresolved at a global level and some ideas that have the potential to come into fruition.

It has been argued that there is no internationally accepted legal definition of 'biological weapon' that goes beyond the BWC's vague definition, which turns on the possessor's intent.[98]In addition, there is no authoritative list of prohibited agents. Therefore, a suggestion is to create a multi-layered approach to defining bioweapons, providing a list of agents that are prohibited in certain quantities, no matter the circumstances.[99] However, a disadvantage of this approach is that it would be impossible to include new bioweapons that have not been invented yet, deeming the list always defective. Therefore, introducing a mechanism or software that would quickly add new weapons to the list once approved as 'prohibited' could seem more effective.

My recommendation to improve preparedness is the creation of simulated scenarios that would consider the integration of multiple modes of management and risk analysis which would testify to different techniques of operations, allowing the estimation of the effectiveness of different defensive responses. Scholars recommend two public health response models, known as the coercive model and the cooperative model. The coercive model relies on aggressive measures such as quarantines, confinements, mandatory vaccinations and other forceful methods to contain the spread of disease.[100] On the other hand, the cooperative model respects civil rights and nurtures public trust in health officials in order to encourage compliance with reasonable emergency measures.[101] Supporters of this model argue that public health officials should provide clear and accurate information to the public in the event of an attack and should structure health responses to be as respectful of individual rights as possible.[102] The creation of stimulated scenarios using these two different responses would provide a realistic assessment of both responses resulting in affecting the allocation of resources during a crisis, the mindset of public health officials and the measures they consider reasonable.[103]

Another recommendation to strengthen preparedness, is the establishment of an 'in-house' committee with scientists, epidemiologists and medical professionals from all nations who would work towards the development and stockpiling of drugs and vaccines. The focus of

[98] Eric Merriam, 'The International Legal Regime Affecting Bioterrorism Prevention' [2014] 3 National Security Law Journal 1.
[99] ibid., 33.
[100] Fouchier and Osterhaus (n 28), 296.
[101] George J. Annas, 'Terrorism and Human Rights' in Jonathon D. Moreno (ed), *In The wake of Terror: Medicine and Morality in a Time of Crisis* (MIT Press 2003) 43.
[102] Fouchier and Osterhaus (n 28).
[103] ibid., 297.

the committee would be the adoption of a proactive ('just-in-case') strategy to develop vaccines against infectious diseases with pandemic potential,[104] not only for pathogens that have been used in the past but search for new biological agents and using the latest biologic and genomic tools, trying to develop new diagnostic methods, therapies, and vaccines.

Another recommendation considers business level regulation. The private sector plays a vital role in the economy and the growing risk of bioterrorism makes the business sector vulnerable to disruption.[105] Therefore, the recommended UN legal body should encourage business' participation in defending against bioterrorist attacks. The UN legal body may issue business level regulations or issue an extended list of guidelines and standard practises (after researching as to what is effective in a business environment and what is not), giving the freedom and flexibility to businesses to determine what measures they will adopt in order to improve preparedness in a way that best fits their business models.[106] However, the adoption of the regulations and improvement of preparedness requires a good amount of investment on behalf of the businesses. The recommended UN legal body should be able to source funding through the nation's contribution and allocate a budget dedicated to helping businesses implement the regulations.

Another recommendation would be introducing a multidisciplinary intelligence service within the UN body responsible for investigating, scrutinising, and overseeing the compliance of the laws and regulations within the nations. This would also solve the problem of verification mentioned in Chapter 2. In doing so, the intelligence service could create a unique software of the highest technology that would allow real-time exchanging, collecting, and analysing data and forensics. The aim of the service would be the absolute attribution of a biological weapon to its source, including the identification of persons, places, processes, or instruments involved in the attack.[107] The existence of the service would have a dual scope, firstly to punish the attackers as their work would lead to the criminal prosecution of those involved through the collection of evidence and secondly, a preventative scope as their work would lead to the elimination or at its best, complete exclusion of the possibilities of a biological attack.

Consistent, clarified, and updated law that prepares for the consequences of a bioterrorist attack, minimising time-consuming lawsuits and confusion over who is in charge in an emergency, is one aspect of a solution for an effective response in case of bioterrorism. However, some of the most difficult tasks that the recommended UN legal body would have to face is to find funding, develop training programs, source solid personnel (since bioterrorism is often overlooked within educational institutions as the focus is more on teaching conventional forms of attacks), convince all nations to become part of this new legal entity (as some nations have their own political agenda and may refuse participation), acquire the sources to conduct continuous research in order to respond to the rapid changes in biotechnology, and

[104] National Research Council (n 13).
[105] US Department of Homeland Security, 'Homeland Security Presidential Directive 7: Critical Infrastructure Identification, Prioritization, and Protection' (*Cybersecurity & Infrastructure Security Agency*, 17 December 2003).
[106] Gigi Kwik Gronvall, *Preparing for Bioterrorism, The Alfred P. Sloan Foundation's Leadership in Biosecurity* (UPMC 2012) 118.
[107] National Research Council (US) Panel on Biological Issues, *Countering Bioterrorism: The Role of Science and Technology* (National Academies Press 2002).

to evaluate the recommended practises to establish their effectiveness, are but a few problems that require a solution.

In conclusion, this article recommends the introduction of a UN body with law enforcement power responsible for overseeing the compliance of its laws under which all countries should collaborate to address the root causes of bioterrorism and develop appropriate preparedness and preventive strategies.[108] As mentioned above, effective preparedness is a deterrent to bioterrorism. However, the law itself is not sufficient to tackle bioterrorism. More training programs provided by educational institutions, improvement of facilities in order to carry out research, development of deradicalization schemes for terrorists convicted of bioterrorist attacks as part of the restorative justice, as well as the development of new technology to detect the transmission of biological agents, such as the introduction of border security checks or screenings for biological weapons are needed in order to be fully prepared for a biological attack. My recommendation regarding establishment of a UN body could act or campaign towards that direction so the above ideas can be achieved.

Conclusion

This article has argued that bioterrorism is one of the greatest threats modern societies face today. There is no criminal threat with greater potential danger to all countries and people in the world than the threat of bioterrorism.[109] This is because the extent of disaster that bioterrorism can cause has no geographic, economic or political boundaries, as an incident in one state is enough for the disaster to reach a global dimension due to the movement of people and goods. Therefore, it requires a unique approach as the first responders to mitigate the threat will be the health professionals and the scientific community, analysing the data and deploying the vaccines.

However, science and technology have played a great role in increasing this threat. As a result, there has been a great debate on whether science should be regulated and to what extent. The scientific community and policymakers need to cooperate in a new way to find the right balance between open science and security, ensuring that the oversight (which in my opinion is needed) does not pose barriers to the evolution of science.

Unfortunately, as discussed in Chapter 2, the lacuna in the current international law means that the law cannot prevent bioterrorism. This is because there are many grey areas in the existing Treaties. As a result of their ambiguity, most of these Treaties have a consultative role and are treated as 'soft law.' In addition, the international law on bioterrorism needs to be modernised as biotechnology is growing rapidly and having 'future proof' treaties are important to follow the changes in technology.

Finally, various recommendations have been proposed in this article concerning core aspects that are relevant to the bioterrorism problem. It has been suggested to give authority to the WHO to oversee research projects and introduce verification and enforcement mechanisms to keep the Treaties alive and avoid greater balkanisation of the law. Furthermore, as illustrated

[108] Chevrier (n 35).
[109] Sandy Smith, 'Interpol Hosts First Global Conference on Bioterrorism' *EHS Today* (24 February 2005) <https://www.ehstoday.com/archive/article/21911884/interpol-hosts-first-global-conference-on-bioterrorism> accessed 17 February 2021.

in the last chapter, this article has recommended the establishment of a new UN legal body which would enforce common policies to all nations at a global level. This prospective UN body would fill the gap of not having an international institution charged with regulating biosecurity globally. It would be responsible for resolving complex issues that concern bioterrorism, such as education of personnel, introducing UN-based intelligent offices and funding preventative studies. Without this legal framework, the aim to prepare for bioterrorism will be lost in a system of politics without solutions, laws without obedience and a future without security.

Changes of use and the planning system: *Sage v Secretary of State for Housing, Local Government and Communities*

Thomas Spencer

Introduction

The pandemic has triggered a revolution in homeworking. Technological change has now made it possible for large numbers of workers to do their job without having to enter an office. However, certain conduct to facilitate working from home could be unlawful. The Town and Country Planning Act (TCPA) 1990 requires that planning permission is obtained where a 'material change of use' occurs to a building.[110] This requirement is limited by the rule that changes that are 'incidental to the use of the dwelling house' do not require planning permission.[111]

In the recent case of *Sage v Secretary of State for Housing, Local Government and Communities*,[112] the High Court clarified the meaning of whether a change in use is 'incidental.' The key test was whether there has been a change in the character of the use. The following shall explain and evaluate Sir Duncan Ousley's judgement and make comments on the efficacy of the requirements of the Town and Country Planning Act 1990.[113] Although the application of law was correct, I will submit that this judgement does reveal critical issues with the English method of land use regulation indicating a need for reform to planning law.

Facts

Mr Sage ('the appellant') is a personal trainer who ran his business out of a timber outbuilding with windows in his back garden. This outbuilding sits 20 metres deep into the garden and is fitted with a treadmill, cross-trainer, free-weights, balls, bench, and punch bag, and is only accessible through a passage shared with the neighbouring property in between their homes. This outbuilding, which was initially used as a shed, has been used as a gym since 2016 to support his work as a personal trainer. However, the appellant failed to apply for planning permission when making this conversion, and thus has twice applied for a Certificate of Lawful Use (CLU) to Bromley London Borough Council (LBC).[114] This certificate would be conclusive as to the lawfulness of the use of the property.

The second application was done to overcome several evidentiary shortcomings identified in the first refusal but was still rejected for the same reasons; chiefly, that this constituted development as defined under section 55 of the TCPA 1990,[115] consequently requiring that planning permission is achieved before the conversion of use can be considered legal. The appellant appealed this decision to the Secretary of State for Housing, Local

[110] Town and Country Planning Act (TCPA) 1990, s 55(1).
[111] TCPA (n 1), s 55(2)(d).
[112] [2021] EWHC 2885 (Admin).
[113] TCPA (n 1).
[114] TCPA (n 1), s 191.
[115] ibid.

Government and Communities ('the respondent') resulting in an inspector conducting a virtual site visit and hearing. The inspector dismissed the appeal on the 10th of February 2021. The appellant disputed the findings of the inspector on several issues. Most prominently that the inspector took an immaterial consideration into account; namely, the visual disturbance presented by the outhouse. He also claimed that any changes were 'incidental' under section 55(2)(d) of the TCPA 1990. This would mean it did not constitute a development, and thus would not require planning permission. Finally, the appellant claimed the inspector's decision was irrational.

The Judgement

The judge, Sir Duncan Ousley, began by addressing the notion that the changes were merely incidental. If it was the case that any material changes were merely 'incidental to the enjoyment of the dwelling house,' then permission would not have been required. This would have been sufficient to allow for the CLU to be given. However, Sir Ousley raises a history of case law presenting a significantly higher burden to finding changes that are incidental than what is presented by the appellant. For example in *Burdle v Secretary of State for the Environment*,[116] Bridge J found that changes cannot be considered incidental where they differ from 'a single main purpose of the occupier's use of his land.'[117] This is supported by a more recent judgement in *Wallington v Secretary of State for Wales*.[118] Here, it was found that even changes to support non-commercial activities are not necessarily incidental if they represented a situation where a dwelling-house is used as anything else but a dwelling house.[119] In this case significant changes were done to support the outbuilding changing from a domestic purpose of storing goods to a commercial one operating as a gym. Sir Ousley interpreted this as meaning that the development was not incidental to the enjoyment of the outbuilding as a shed.

The second argument that the judge addressed is related to the alleged immateriality of the consideration of environmental harm resulting from the development. The inspector found that there was a visual disturbance from the frequent entry and exit to the garden in a way that was greater than what would be normal in a purely domestic garden. This was problematic, in the view of the Inspector, because the garden was visible from the window of the neighbouring property. The Secretary of State's Planning Practice Guidance clearly states that pollution is a factor the decision maker may consider; this is not limited to air or noise, but also visual pollutants. Moreover, the increased foot-flow into the garden would have more noise when gym goers are entering and leaving the outhouse.

The question of legal inadequacy was also answered. The claimants had relied upon *South Buckinghamshire District Council v Porter*[120] to argue that the inspector had not sufficiently well explained the decision to reject the CLU based on the visual disturbance. However, this argument, in the view of the judge, mistakes the level of detail required in the letter. The decision-maker must merely make it clear what their reasoning was, as opposed to show why their observation is persuasive. The inspector detailed quite clearly how on an

[116] [1972] 1 WLR 1027.
[117] ibid, [5].
[118] [1991] 62 P&CR 150.
[119] ibid, [6].
[120] [2004] UKHL 33, [36].

average day, there will be as many as five additional customers coming in and out of the outhouse than what would be expected in a normal garden thus presenting, in their view, a visual disturbance.

Finally, the irrationality argument was answered; namely, that the claimant had submitted that the finding of any disturbance from the change of use was irrational. Given extensive evidence of there not being any noise pollution, the claimants felt that the only disturbance could be visual, and mere visual disturbance from one window of less than a dozen clients a day could not rationally constitute a disturbance in the eyes of the claimants. However, there was no evidence that noise pollution did not exist since permission was not granted to adduce witness statements. Moreover, the test the inspector was applying was not whether planning permission should be refused; rather, it was whether permission was required. These considerations are useful in detailing that this change constituted development under the TCPA 1990,[121] and thus cannot be seen to be irrational.

Comment and Reform

The key takeaway from this case is that the application of law is completely correct. Whilst it still may be possible for permission to be obtained for a change of us, the claimant has failed to present sufficient evidence that the changes he has made are lawful. However, the interesting thing about this is from a public policy ground it is very difficult to see why such an action should be unlawful. There was absolutely no noise pollution due to windows and air conditioning that meant the windows did not need to be opened. There were no visual changes to the outside of the outhouse meaning there's clearly no visual pollution other than the entry and exit of customers into the gym. Moreover, neighbours had commented that they didn't mind or object to the change. Given all of this, it is very hard to see who the victim was, and thus what the need is for regulation prohibiting such changes.

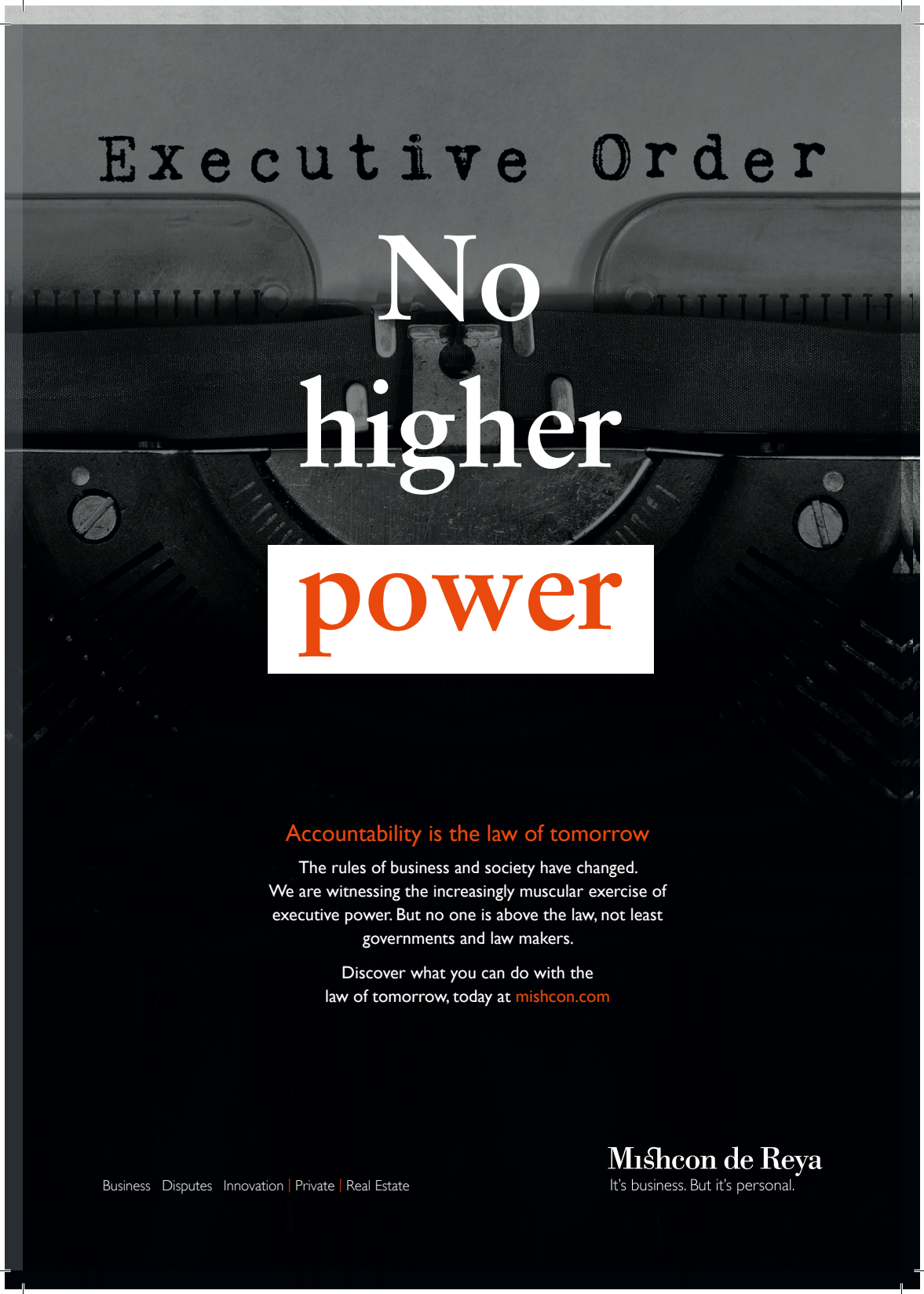
Developments do create negative externalities,[122] and for this reason, we require rules regulating which ones are permitted and which ones aren't. However, in this case the rules have prohibited an entrepreneurial individual from providing a service demanded by enough customers to constitute a business. As a result of the refusal to grant a CLU, it could become the case that this business is no longer plausible, a situation that benefits no one. This represents a major flaw in our planning law. Given the clear relationship between constraints on development and economically damaging outcomes,[123] we should aim to avoid unnecessary interventions in development rights where possible. Clearly, due to the lack of objections and pollution from the change of use, no one is harmed by this change, and it benefits both the owner and his customers tremendously. A good way around this would be an amendment to the TCPA 1990 where it is made possible for a CLU to be obtained where a sufficient number of supportive letters from neighbouring residents have been written. This combines the public

[121] TCPA (n 1).
[122] Esteban Rossi-Hansberg, Pierre-Daniel Satre and Raymond Owens III, 'Housing Externalities' [2010] 118 Journal of Political Economy 3.
[123] Christian Hilber and Wouter Vermeulen, 'The Impact of Supply Constraints on House Prices in England' [2016] 126 The Economic Journal 591.

demand for community involvement in the planning system without unnecessarily preventing beneficial developments like Mr. Sage's change of use.

Conclusion

From a legal point of view the application of law in this case is orthodox and correct. Based on the current way CLUs are understood, the appellant could not show that they had met the criteria to do so. Yet from a policy perspective there is no good reason as to why Mr. Sage should not be allowed to continue legally running his business. There are no victims, and all his clients and himself benefitted from the services he provided. This is indicative of a law rife for reform, and thus, it is submitted that the Government in their promises for planning reform look to better ways to consider CLU applications so that developments clearly in the public interest can be permitted.



No higher power

Accountability is the law of tomorrow

The rules of business and society have changed. We are witnessing the increasingly muscular exercise of executive power. But no one is above the law, not least governments and law makers.

Discover what you can do with the law of tomorrow, today at mishcon.com

Business Disputes Innovation | Private | Real Estate

Mishcon de Reya
It's business. But it's personal.

Practical Aspects of Application of the Rules of *lex mercatoria* in International Commercial Arbitration

Daria Efimova

The most important factor in determining applicable law in international commercial arbitration is the reasonable expectations of the parties. If parties have stipulated that a dispute shall be resolved in accordance with the established rules of law, an arbitrator must apply them.[1] However, a very complex issue in international commercial arbitration is determining the applicable law when the parties make no choice. In the absence of express agreement, it is generally considered that an arbitrator must proceed in three steps. Firstly, asking himself whether the application of a national system is appropriate. Then, if not, whether he/she should apply national rules. Finally, in the latter circumstance, existing national rules are considered in relation to relevance to the dispute. One aspect of this requirement is the neutrality of substantive law. Thus, there are several factors that influence the determination of applicable law.

The first set of factors concerns the parties themselves. The fact that they have different nationalities is important and possibly irreplaceable. Moreover, one of the parties can be transnational, because such entities should be seen as existing in a free space detached from any nationality. Applying the national law of one of the parties by arbitrators in the absence of a choice-of-law clause can be seen as biased. Other indicators relate to the nature of transactions. An example would be a contract between individuals of the same nationality to transport goods from one country to another or for the provision of services abroad.

Lex mercatoria, as a neutral and alternative regulator of international commercial relations, was designed to address these difficulties. Currently there is no single definition of *lex mercatoria* that fits all opinions about the nature of this phenomenon. Firstly, *lex mercatoria* is determined as an independent legal order, which includes the law of international trade, which is spontaneously created by parties during international economic transactions.[2] Thus, the parties to international commercial relations find themselves outside the legal field of a single country, regardless of whether they have chosen it as applicable to the contract or not. Secondly, *lex mercatoria* is perceived as a set of rules that are exhaustive for the resolution of a commercial dispute. Under this concept, *lex mercatoria* exists as an alternative to the applicable law.[3] Thirdly, *lex mercatoria* is a consolidation of customs and basic principles of international trade conventions.[4] It follows from this position that *lex mercatoria* does not exclude the application of national law, and in some cases may even form part of it. By subjecting a contract to rules common to all legal systems, they may also subject it to the rules

[1] Julian Lew, 'Applicable Law in International Commercial Arbitration' (Oceana/Publication, 1978) 75.
[2] B Goldman, 'Frontières du droit et lex mercatoria' (9 Archives de philosophie du droit, 1964) 177.
[3] Editors, *Jowitt's Dictionary of English Law*, 'Lex mercatoria' (5th edn); Editors, *Osborn's Concise Law Dictionary*, 'Lex mercatoria' (12th edn).
[4] Lowenfeld Andreas, 'Lex Mercatoria: An Arbitrator's View' (Arb Int'l, 1990) 133.

of the state involved in the dispute. If such general rules cannot be determined, the arbitrator will apply the rules that seem fair and most appropriate. Some scholars also believe that *lex mercatoria* is a quasi-legal recognition of the rules of common sense and fairness. In this case, *lex mercatoria* is viewed as principles that do not form the applicable law.[5] Gunter Teubner, in turn, suggested a unique approach.[6] Using a biological term, he argues that *lex mercatoria* is an autopoietic system. Autopoetic means automatic. In Prof. Teubner's reflection, *lex mercatoria* is not a set of material norms, but rather a process in which it organises and produces these norms itself.

In addition to doctrinal definitions of *lex mercatoria*, one can find attempts to define this phenomenon in judicial practice. Judges of the Italian Court of Cassation once stated that a law under which arbitration operates independently of the laws of various states must be regarded as transnational law.[7] 'Mercantile' law (*lex mercatoria*) exists through the business community's adherence to the values of its environment, which represent *opinio necessitatis*, i.e., a dominant belief regarding their binding nature.

Despite the lack of a universal definition, there are a number of provisions that reflect most opinions concerning the nature of *lex mercatoria*. These are, first of all, the exclusion of the law of international trade from any particular legal system, whether international or national. The use of *lex mercatoria* also abandons traditional conflict of laws principles as the main tool of private international law and, finally, includes non-legal provisions, such as model forms of contracts.

The analysis of case practice shows several conditions under which the application of the rules of *lex mercatoria* becomes possible. Firstly, when the choice of *lex mercatoria* has explicitly been made. As the official commentary on the preamble to the UNIDROIT Principles of International Commercial Contracts (hereinafter – the UNIDROIT Principles) notes, there is an increasing tendency to permit the parties to agree on 'rules of law' that are not rules of any national system.[8] In this case, the UNIDROIT Principles will apply to exclude any national law without affecting only super-injurious legal provisions.[9] Similar regulation can be found in Art. 28 (1) of the 1985 UNCITRAL Model Law on International Commercial Arbitration (with amendments as adopted in 2006) and Art. 42 (1) of the 1965 Washington Convention.[10] In more precise terms, the XIII.4.1 'Trans-Lex' Principle, which is called 'Rules applicable to merits; decision *ex aequo et bono*' states that 'the arbitral tribunal shall decide the dispute in accordance with such rules of law as are chosen by the parties as applicable to the substance of the dispute.'[11] In the ruling of the International Commercial Arbitration Court at the Chamber

[5] Highet Keith, 'The Enigma of the Lex Mercatoria' [1989] 63 Tulane Law Review 627.
[6] Teubner G, 'Global Bukowina: Legal Pluralism in the World Society' (Global Law Without a State, 1997) 12.
[7] Case No 2285 of 8 February 1982, Court of Cassation.
[8] Stefan Vogenauer, *Commentary on the UNIDROIT Principles of International Commercial Contracts (PICC)* (2nd edn, Oxford 2015) 37; Bonnell Michael Joachim, 'UNIDROIT Principles: A Significant Recognition by a United States District Court' (Uniform Law Review vol IV, 1999) 207.
[9] Kanashevsky V, 'The concept of 'lex mercatoria' in international private law' (Russian Yearbook of International Law, 2008) 106; UNIDROIT Principles of International Commercial Contracts (PICC), Art 1.4.
[10] Convention on the Settlement of Investment Disputes Between States and Nationals of Other States - International Centre for Settlement of Investment Disputes, Washington 1965.
[11] TransLex-Principles, Art XIII.4.1 <http://translex.uni-koeln.de/principles/of-transnational-law-(lex-mercatoria)> accessed 20 November 2021.

of Commerce and Industry of the Russian Federation (hereinafter – the ICAC at the RF CCI), the parties contractually stipulated that all disputes, controversies and claims arising out of or in connection with the contract, including those related to its interpretation, performance, breach, termination or invalidity, shall be settled based on general principles of law (*lex mercatoria*) before the ICAC at the RF CCI, Moscow, in accordance with its Rules of Procedure.[12] The arbitral tribunal held that under Art. 28 of the Law of the Russian Federation 'On International Commercial Arbitration', which is based on the UNCITRAL Model Law, in rendering its award, the ICAC at the RF CCI is guided by the rules of law which the parties have chosen as applicable to the merits of the dispute.[13] The parties stipulated in their contract that disputes from the contract were to be resolved on the basis of general principles of *lex mercatoria* and that all conditions not regulated in the said contract shall be governed by the laws of Germany and the Russian Federation. The arbitral tribunal decided to apply the UNIDROIT Principles and terms of the contract as the parties had not chosen the law of a particular state. Therefore, arbitrators granted authority to the parties' explicit choice of *lex mercatoria* as the law applicable to the contract.

However, in practice, when parties to an international commercial contract wish to avoid applying national laws in the event of a dispute, they rarely include a clause with the term *lex mercatoria* in their contract. More often they enshrine a clause referring to the customs of international trade, general principles of law, transnational law, etc. Such contractual clauses also allow arbitral tribunals to apply the rules of *lex mercatoria.*[14] In the well-known 1977 case between the Libyan Arab Republic and the Libyan American Oil Company (LIAMCO), the concession agreement contained a proviso that the agreement shall be governed by and construed in accordance with 'the general principles of Libyan and international law' and, in the absence of such general principles, in accordance with 'the general principles of law.'[15] It was thus established that the primary law governing the agreement was Libyan domestic law, unless it contradicted the principles of international law, and that any part of Libyan law contrary to the principles of international law had to be excluded. In the course of the case, the tribunal concluded that reference should be made to the sources that have been adopted by the International Court of Justice to determine the meaning of 'principles of international law.' 'Principles of Libyan law' can be drawn from Libyan and Islamic law. As to 'general principles of law', the tribunal stated that they are generally embodied in the most recognized legal systems and, in particular, Libyan law, including its modern codes and Islamic law. The court concluded that the general principles of law constitute a body of legal precepts and rules of conduct generally accepted in theory and in practice. Examples include the principle of inviolability of property, the prohibition of unjust enrichment, the obligation to compensate in the event of wrongful damage, etc.

If the parties choose arbitration to settle their dispute without choosing a specific law governing the essential issues of the dispute, then, as happened with merchants in the Middle

[12] Case 11/2002 [2002] ICAC at the RF CCI.
[13] Law of the Russian Federation 'On International Commercial Arbitration' no 5338-1 (as amended on December 29, 2015), Art 28.
[14] International Institute for the Unification of Private Law Principles (UNIDROIT Principles) 2016, Preamble.
[15] *LIAMCO v The Government of the Libyan Arab Republic*, YCA 1981.

Ages, the commercial court, i.e., arbitrators will apply the general rules of the commercial community. In the *Saudi Arabia v Aramco* decision, arbitrators referred to the rules of *lex mercatoria* to fill the gaps in Saudi law.[16] In *Sapphire International Petroleums Ltd v National Iranian Oil Company,* an arbitrator applied the rules of *lex mercatoria* instead of national law, arguing that national law can be changed by the state and is often unknown or not well known to one of the parties.[17] In the case of *Norsolor S.A. v Pabalk Ticaret Limited Sirketi,* an arbitrator chose directly *lex mercatoria* to decide the dispute.[18] This decision was later challenged in Austria and France, but courts found that an arbitrator had not exceeded his powers by choosing *lex mercatoria* and upheld the decision[19]. To exclude the application of *lex mercatoria* in these institutions would deprive merchants of the last opportunity to settle disputes according to their needs. It can be suggested that the absence of a choice-of-law clause is an indication that the parties do not want to apply any national system. The rationale behind this statement is that the parties could not agree on which systems should govern their relations. The parties, for instance, could not possibly agree on the choice of applicable law because each party wishes to choose its 'home' law, or the purported applicable law is so undeveloped that it is incapable of providing a solution to the issue at hand.[20] However, this suggestion ignores the possibility that the choice of national law was so obvious that it was not worth mentioning or that the parties had never thought about the issue at all. In resolving the *Mechema Ltd. v S.A. Mines* dispute, the tribunal concluded that it was impossible to determine the law of one party as applicable from the text of the contract and its substance, so arbitrators referred to *lex mercatoria* in their reasoning.[21] This, however, does not prevent the parties from excluding *lex mercatoria* as the law governing the dispute. If such exclusion is not contained in an arbitration clause or arbitration agreement, the rules of *lex mercatoria* apply. If the rules of *lex mercatoria* prove insufficient to resolve the dispute, the remaining questions may be answered by national law, which conflict-of-law rules will determine.

When addressing the question of conflict-of-laws, the answer can be found in international conventions and procedural rules established by arbitration institutions. There are several approaches to dealing with the conflict-of-laws rules. The first one is to apply the rules that arbitrators consider to be applicable. The European Convention on International Commercial Arbitration provides in Art. VII that 'failing any indication by the parties as to the applicable law, the arbitrators shall apply the proper law under the rule of conflict that the arbitrators deem applicable.'[22] Similar wording is contained in the 2021 Arbitration Rules of

[16] *Saudi Arabia v Arabian American Oil Co* (ARAMCO) [1963] ILR.
[17] *Sapphire International Petroleums ltd v National Iranian Oil Company*) [1963] ILR; Matti Kurkela, *Letters of Credit Under International Trade Law: UCC, UCP and Law Merchant* (Oceana Publications,1985) 14, 331.
[18] *Pabalk Ticaret Limited Sirketi v Norsolor*, ICC Case no 3131; Emmanuel Gaillard, 'France: Court of cassation decision in Pabalk Ticaret v Norsolor (Enforcement of Arbitral Awards; Lex Mercatoria)' (International Legal Materials vol 24 no 2, 1985) 360.
[19] Judgement of the French Court of Cassation [1984] 83-11.355; Decision of the Austrian Supreme Court [1982] 8 Ob 520/82.
[20] Marco Mastracci, *The International Distribution Agreement: A Practical Approach to Transnational Contracting across the European Union, the United States and Latin America* (2nd edition, 2020) 26; Case 7710/1995 [2001] ICC.
[21] *Mechema Ltd v SA Mines Minérais et Métaux*) [1982] Belgium VII Y.B. Commercial Arbitration 80.
[22] European Convention on International Commercial Arbitration 1961, Art VII.

the International Chamber of Commerce (hereinafter – the ICC).[23] Article 21 provides that 'in the absence of any such agreement, the arbitral tribunal shall apply the rules of law which it determines to be appropriate.' This approach can also be found in the 1985 UNCITRAL Model Law on International Commercial Arbitration. In this sense, the achievement of uniformity is facilitated by UNCITRAL's work to publish the generalised practice of national courts and international arbitrations.[24]

Case practice has long followed this position. In *ICC Case No. 9419*, the parties had failed to establish the applicable law in their contract. The claimant argued that, in the absence of any reference to a specific national law, the claims were based on generally recognised principles of international trade, allowing the arbitrator to base his award on the provisions of *lex mercatoria*, whose codified rules can be found in the UNIDROIT Principles. The arbitrator, in turn, pointed out in the award that, in the absence of an express choice of law by the parties, the arbitrator could not choose the applicable law directly but had to rely on the conflict-of-laws rules that he considered to be the most appropriate.[25]

The second approach is to use the rules of the seat of arbitration (*lex arbitri*) in the absence of an agreement on applicable law by the parties.[26] However, this solution has often led to undesirable results because the parties have chosen the place of arbitration for practical reasons rather than based on local conflict-of-laws rules.[27]

The final approach in determining the applicable law is for the arbitral tribunal to decide the dispute because of those rules which are most closely related to the dispute, including *lex mercatoria*. In this case, arbitral tribunals abandon the conflict-of-laws method.[28] *The Resolution of the Plenum of the Supreme Court of the Russian Federation on International Private Law No. 24* identifies the criteria applying the 'closest connection' principle in much detail.[29] Apart from the most used ones such as place of residence, establishment or main activity and place of performance, additional criteria, which may be considered, are introduced. For example, courts are invited to determine which country's law will best implement generally recognised principles of civil law and the construction of certain civil law institutions. The Supreme Court also recalls that such principles and institutions include the prohibition of taking advantage of bad faith conduct, the prohibition of abuse of right, the protection of a weak party, the preference to preserve the validity of the transaction, and the prohibition of unreasonable refusal to perform an obligation.

[23] International Chamber of Commerce Arbitration Rules 2021, Art 21 <https://iccwbo.org/dispute-resolution-services/arbitration/rules-of-arbitration/> accessed 16 November 2021.
[24] United Nations Commission on International Trade Law Editors, *UNCITRAL Digest of Case Law on the United Nations Convention on the International Sales of Goods* (United Nations 2016).
[25] Case 9419 [1998] ICC.
[26] William W Park, 'The Lex Arbitri and International Commercial Arbitration' (International and Comparative Law Quarterly vol 32 no 1, 1983) 22-24.
[27] Frank Baddack, *Lex Mercatoria: Scope and Application of the Law Merchant in Arbitration* (University of the Western Cape 2005) 5 <https://etd.uwc.ac.za/bitstream/handle/11394/250/Baddack_LLM_2005.pdf?sequence=1&isAllowed=y> accessed 24 November 2021.
[28] Trans-lex.org 'XIII.4.1 - Rules applicable to merits; decision ex aequo et bono' http://www.trans-lex.org/970020> accessed 21 November 2021.
[29] Resolution of the Plenum of the Supreme Court of the Russian Federation on International Private Law no 24, 'On the Application of Rules of International Private Law by the Courts of the Russian Federation', Art 6.

In 2008, the ICC International Court of Arbitration heard a case between two companies, established in country A and country B, which entered into an agreement for the provision of high technology services. The agreement was silent on the law applicable to the contract. It only stated that, in the event of a dispute, the parties would go to the London Court of International Arbitration (LCIA), and the proceedings would be conducted in English. The dispute arose when one company accused the other of failing to fulfil its obligations. The claimant brought the case before the ICC International Court of Arbitration, and the respondent did not object. The claimant argued that the tribunal should have applied the English law because London had been chosen as the seat of arbitration in the agreement, English had been chosen as the language of the arbitration and English law was neutral. The respondent objected to the claimant's position stating that the national law of country B was applicable. The arbitral tribunal held in the interim award that there had been no express agreement between the parties as to the choice of law applicable to the case. It decided to choose the UNIDROIT Principles as the applicable law and, if necessary, apply the law that is considered appropriate.

When arbitrators finally conclude that *lex mercatoria* is the applicable law to the dispute, the content of the rules of *lex mercatoria* will be established as follows. Firstly, arbitrators may choose to use one of the documents that are considered to be an authoritative codification of principles of *lex mercatoria* ('registry method').[30] In practice, in addition to the UNIDROIT Principles referred to above, the Principles of European Contract Law (Principles, Definitions and Model Rules of European Private Law) are also widely used.[31] The Trans-Lex Principles, developed by the Transnational Law Centre headed by Klaus Peter Berger, may also be added to this list. However, Mr. Berger, who defends the principle of 'progressive codification' of *lex mercatoria*, criticises the 'registry method for the lack of flexibility in light of the speed of development of modern law and business practices.[32] Secondly, arbitrators may pose specific questions in connection with a disputed legal relationship, to which they will seek answers in a variety of sources ('functional method').[33] Arbitrators do not limit themselves by relying on a certain 'register' of principles of *lex mercatoria* while using the 'functional method.'[34] It seems reasonable to agree that 'functional method' better reflects the legal nature of *lex mercatoria* as its sources are not limited to unification acts. The general principles of law and standard forms of contracts also have a great influence on the trade turnover.

The analysis of the aforementioned cases draws a general conclusion that the role of *lex mercatoria* in the regulation of foreign trade relations is becoming increasingly important despite the ambiguity of its legal nature. This is confirmed by the weight of authority of already existing documents and the expansion of their scope. By applying the rules of *lex mercatoria*

[30] Nigel Blackaby and others, *Redfern and Hunter on International Arbitration* (6th edn. Kluwer Law International; Oxford University Press 2015) 219.
[31] Principles of European Contract Law – PECL 2002, Parts I, II; Study Group on a European Civil Code and the Research Group on EC private law (Acquis Group), *Principles, Definitions and Model Rules of European Private Law: Draft Common Frame of Reference – DCFR* (Sellier European Law Publishers 2008) 412.
[32] Luke Nottage, ed Dr Klaus Peter Berger, 'The practice of transnational law' (Kluwer Law International, 2001) 371.
[33] Blackaby Nigel, Constantine Partasides et al., *Redfern and Hunter on International Arbitration* (6th edn. Kluwer Law International; Oxford University Press 2015) 219.
[34] Loboda Andrey, 'The controversy surrounding the modern law of merchants ('lex mercatoria')' (MGIMO University) <https://mgimo.ru/upload/docs/Slide3.pdf> accessed 24 November 2021, 49.

the parties avoid rules that are not appropriate for international contracts, such as strict formalities. In addition, neither party will have the advantage of regulating the dispute by its own law. These are the reasons why despite the recommendatory nature of the rules of *lex mercatoria*, it has been widely recognised both in doctrine and case practice that the parties are free to choose not only the law of a particular country to be applicable to the contractual relations but also rules of a non-national character.

Sacrificing Legal Certainty on the Altar of Market Integration: The Difficult Road Towards an Internal Market in Goods

Richard Ninov

Introduction

Free trade is, as European Union (EU) values go, decidedly uncontroversial. Though the desirability of free movement of persons and services within the internal market has been questioned by some, unfettered commerce is unanimously embraced.[35] Given this appeal, whatever the future of the internal market, the free movement of goods will perhaps prove to be the most resilient of the 'Four Freedoms.' However, the Union's attempts at integration in this area have been marred by what seems to be a wavering of conviction among the Courts. This work will consider the EU's attempt at negative integration on the internal market in goods by briefly examining the law surrounding both tariff barriers and non-tariff barriers. It will be argued that the jurisprudence surrounding the latter, in particular that on measures having equivalent effect to quantitative restrictions, has developed in an unsatisfactory manner – a phenomenon invited by an apprehensive judicial interpretation of Articles 34 and 35 of the Treaty on the Functioning of the European Union (TFEU).[36] Finally, it will be emphasised that this has not only been inconsistent with the EU's objectives, but also detrimental to legal certainty in this area.

Tariff and Non-Tariff Barriers – A Double Standard?

The establishment of an internal market in goods has been achieved through both positive and negative harmonisation. The former, also referred to as 'approximation' of EU law, aims to close the gap between national regulatory regimes by laying down positive rules through directives.[37] By contrast, negative harmonisation concerns the striking out of national rules which constitute 'obstacles to intra-Community trade', as recognised in *Gaston Schul*.[38] Both of these integrative tools are meant to contribute to the ultimate goal of ensuring a market 'without internal frontiers', as expressed in Article 26(2) TFEU.[39] These can take many shapes, falling into either of two categories – tariff and non-tariff barriers.

Tariff barriers include customs duties and any other charges on foreign goods imposed *at the point of entry* into a Member State. These are comfortably dealt with under Articles 28-

[35] For a Eurosceptic perspective, see MoneyWeek, 'Dan Hannan: what we need from Europe is free trade, not a common government' (3 March 2016) <https://www.youtube.com/watch?v=eMwvWHLRrpE> accessed 5 April 2021.
[36] Consolidated Version of the Treaty on the Functioning of the European [2012] OJ C326/47 (Treaty on the Functioning of the European Union), arts 34-35.
[37] The European Parliamentary Research Service, 'The EU as a community of law' (Briefing, March 2017) 3, 4 <https://www.europarl.europa.eu/RegData/etudes/BRIE/2017/599364/EPRS_BRI(2017)599364_EN.pdf> accessed 5 April 2018.
[38] Case 15/81 *Gaston Schul Douane Expediteur BV v Inspecteur der Invoerrechten en Accijnzen. Roosendal* [1982] ECR 1409, [33].
[39] Treaty on the Functioning of the European Union, art 26(2).

30 TFEU,[40] which prohibit the imposition of any such charges, fiscal or otherwise, on goods arriving from another Member State – a clear example of a negative integrative measure. The Court of Justice of the European Union (CJEU) has likewise been quick to emphasise the importance of unhindered access of the Member States to each other's markets, as seen in *Commission v Italy*.[41] Here, the European Court of Justice (ECJ) accentuated the fact that the *effect* of the charge levied alone is of import, its *purpose* being immaterial. As long as a monetary burden was imposed, it would have the effect of hindering trade. Such a charge would always fall within the meaning of Article 30 TFEU,[42] meaning it would be invalidated.[43]

The second type of tariff barrier is the discriminatory taxation of goods. While customs duties and equivalent charges are levied at the point of entry, discriminatory taxes are a tariff barrier which is charged once the goods are already within the country. Articles 110-113 TFEU treat this type of tariff barrier with the same strictness as the Treaty provisions on customs duties.[44] Logically, the aim is to prevent the prohibitions on customs duties within the internal market from being subverted by simply applying a tariff barrier *after* the goods have entered the country.[45] The approach of the CJEU was thus equally focused on effect – in *Humblot*,[46] the ECJ annulled a tax levied on cars above a certain power rating (an objective and non-discriminatory criterion) simply because it happened to have a disproportionate *effect* on imported cars.[47]

At this point, a trend emerges – the CJEU has led the internal market integration agenda through negative harmonisation. It has consistently struck out tariff barriers and recognised minimal derogations by interpreting the Treaty provisions prohibiting customs duties and discriminatory taxes very widely. To achieve this, the Court has focused on a charge or tax's *effect* on inter-State trade, rather than prioritising the finding of some discriminatory intent to annul it. However, the Court broke this orthodoxy in its approach to non-tariff barriers.

Non-tariff barriers include any measures which restrict the flow of goods through means other than a pecuniary charge on imports or exports. As far as the internal market is concerned, these have been a source of contention and uncertainty. They are dealt with in Articles 34-37 TFEU,[48] qualifying them as either 'quantitative restrictions' or 'measures having equivalent effect to quantitative restrictions' (MEQRs). The application of Articles 34 and 35 TFEU to traditional quantitative restrictions such as quotas or import/export restrictions is simple, as they can be easily identified.[49] There is no interpretive burden in applying the

[40] ibid, arts 28-30.
[41] Case 7/68 *Commission of the European Communities v Italian Republic* [1968] ECR 423.
[42] Treaty on the Functioning of the European Union, art 30.
[43] Though note the sensible exception in Case 18/87 *Commission of the European Communities v Federal Republic of Germany* [1988] ECR 5247, where the ECJ held charges would survive the application of Articles 28-30 where they were levied *only* so as to cover the costs of an EU-mandated inspection of the goods in question, *and* on the condition that they be non-discriminatory in nature.
[44] Treaty on the Functioning of the European Union, arts 110-113.
[45] Craig and De Búrca, *EU Law: Texts, Cases, and Materials* (7th edn, OUP 2020) 706.
[46] Case 112/84 *Michel Humblot v Directeur des Services Fiscaux* [1985] ECR 1367.
[47] Though the strictness of this interpretation may have been somewhat tempered. See Case C-132/88 *Commission of the European Communities v Hellenic Republic* [1990] ECR 1567, where a cylinder-based car tax with similar effect was permitted where its goal was environmental protection – an EU-endorsed policy.
[48] Treaty on the Functioning of the European Union, arts 34-37.
[49] ibid, arts 34, 35.

Treaty provisions in such cases – the prohibition applies in much the same way as with tariff barriers.

Contrastingly, the interpretation of what exactly amounts to an MEQR has been particularly troublesome, and the CJEU's construction of the concept has been inconsistent. The following section examines these developments.

The Evolution of the MEGRs Concept and the *Keck* Blunder

MEQRs have been a difficult matter for the CJEU to address. Whereas the Treaty provisions governing tariff barriers appear clear in actuality, Articles 34 and 35 provide little explanation for exactly which Member State policies might amount to MEQRs.[50] The task of deciding how wide or narrow the concept is thus falls to the Court.

The ECJ's decision in *Dassonville* illustrates an initial approach which was consistent with that adopted for tariff barriers,[51] whereby a generous interpretation of the relevant provisions cast a wide net of liability on Member States. Any national rules which were 'capable of hindering, directly or indirectly, actually or potentially, intra-Community trade [were] to be considered as measures having an effect equivalent to quantitative restrictions.'[52] This view of MEQRs survived for several years, clarified by the ECJ's follow-on decision in *Cassis de Dijon*.[53] Any national measure which had the effect of making inter-State trade more difficult would amount to an MEQR, even if it applied indistinctly to both foreign and domestic goods (meaning discrimination was not a *necessity* to declaring a rule incompatible with Article 34 or 35).[54]

This state of affairs gave the CJEU pause. The *Dassonville*[55] judgement set out a wide application of Articles 34 and 35.[56] This was compounded by the *Cassis de Dijon* decision's leaning towards a 'home country' model of integration, which entrenched the principle of mutual recognition in the internal market – no goods 'lawfully produced and marketed in one of the member states' may be subject to a national prohibition within the internal market.[57] This new reality gave rise to concerns about the emergence of a deregulatory spiral, with each Member State attempting to undercut the other by lowering product standards in an attempt to make their own exports within the internal market more competitive. This was also an interpretation which took regulatory autonomy *vis-à-vis* imports and exports away from Member States to a substantial extent.

These pressures led the CJEU to sound a retreat the next time a national rule was challenged. In *Keck*,[58] a French anti-dumping law survived the application of Article 34 by the

[50] ibid.
[51] Case 8/74 *Procureur du Roi v Benoît and Gustave Dassonville* [1974] ECR 837.
[52] ibid 852.
[53] Case 120/78 *Rewe-Zentral AG v Bundesmonopolverwaltung für Branntwein (Cassis de Dijon)* [1979] ECR 649. Note that this decision also establishes possible grounds for derogation from Articles 34 and 35 where the rule is proportionate and for the benefit of public health and safety and commercial fairness, among others. These are likely *additional* to the main grounds for derogation under Article 36. See Craig and De Búrca (n 6) 762-763.
[54] Treaty on the Functioning of the European Union, arts 34, 35.
[55] *Dassonville* (n 17).
[56] Treaty on the Functioning of the European Union, art 34, 35.
[57] *Cassis de Dijon* (n 19) 664, [14].
[58] Cases C–267 and 268/91 *Criminal Proceedings against Keck and Mithouard* [1993] ECR I-6097.

ECJ,[59] despite this having the *effect* of constraining 'actually or potentially' the import of goods from Member States which had no such anti-dumping laws. This was justified by the Court by drawing a distinction between national rules amounting to 'product requirements', and national regulations forbidding certain 'selling arrangements.' The former would contravene the Treaty provisions, whereas the latter would survive (as long as they were not overtly discriminatory).

On one level, *Keck* saw the CJEU concede to the Member States some degree of regulatory autonomy.[60] To achieve that, it chose to fabricate an arbitrary distinction between product requirements and selling arrangements, with no objective guidance on distinguishing between the two. Worse still, no clear reason was enunciated for the making of such a distinction in terms of its impact on the internal market. Both product requirements and selling arrangements can have the *effect* of obstructing the internal market. Going by the letter of the Treaty and the jurisprudence surrounding tariff and non-tariff barriers, *effect* must be the deciding factor. Notably, this concession of regulatory autonomy was made at the expense of market integration, in that it allowed national rules to survive where they could in fact hinder inter-State trade.

On the surface, the Court expressed concern that traders increasingly invoked Article 34[61] to eliminate rules which impeded their commercial freedom.[62] However, it is suggested that the deeper motive behind the decision was the Court's wariness of the political consequences which could flow from a loss of national sovereignty – something which the home country harmonisation model is usually associated with.[63] The Court in *Keck*[64] tried to strike a balance between two irreconcilable concepts – that of market integration on the one hand, and the idea of Member State regulatory autonomy on the other. In a way, the case exemplifies the challenge the EU has faced since its inception; namely, that of fostering an 'ever-closer Union'[65] whilst grappling with its Members' insistence on retaining substantial self-governance. However, the unpalatable truth is that this balancing exercise is one made in futility. Integration is properly viewed as a zero-sum game – it must either be absolute, or it must be abandoned.

Building on this view, it is suggested that the Court made a mistake in shying away from integration when it had the opportunity to solidify it, in what can only be seen as an appeasement of the individual Member States – an assurance that their national sovereignty could survive market integration. However, as long as 'actual or potential' barriers to trade exist,[66] the internal market cannot truly be said to be integrated. This view is consistent with the orthodox approach to tariff barriers, where *any* rules resulting in hindrances to trade are swiftly struck down. Infringement flows from causation, not fault. In the alternative, commercial actors facing an impaired flow of goods between Member States would be forced to consider ways of satisfying *national* market, rather than *internal* market, rules. One would

[59] Treaty on the Functioning of the European Union, art 34.
[60] *Keck* (n 24).
[61] Treaty on the Functioning of the European Union, art 34.
[62] Ibid 6131, [14].
[63] Barnard and Peers, '*European Union Law*' (3rd edn, OUP 2020) 336-337.
[64] *Keck* (n 24).
[65] A central objective of the European community which has been consistently restated in key instruments, including the Treaty of Rome, the Maastricht Treaty, and the Lisbon Treaty.
[66] *Dassonville* (n 17), 852.

be forgiven for questioning whether such an arrangement deserves to be labelled a 'single market' at all.

Of course, the wide reading of Articles 34 and 35 and the home country model are not without criticism.[67] Resorting to deregulation in such circumstances would be the logical response for each sovereign legislature, though this is something falling outside the Court's remit. Instead, it is the duty of the EU legislators to lay down positive harmonising measures establishing common regulatory standards governing the flow of goods within the internal market. To an extent, centralised positive harmonisation averts the demise of national sovereignty by quasi-democratically allowing individual Member States to participate in the integration process by virtue of their representation in the European Parliament and Council of Ministers.[68]

Conclusion – A Return to Orthodoxy?

For the internal market to function, there can be no room for national protectionism *between* Member States. The integration of the European economies is implicitly conditioned upon a concession of national sovereignty. This means that, for the Union to function sustainably, the regulatory initiative will have to increasingly move *away* from the Member States, and *towards* the European institutions.

The *Keck* decision was inconsistent with the established approach towards trade barriers, which was one based on *effect*.[69] This resulted not only in an undue setback to the European integration project, but also gave rise to considerable legal uncertainty in terms of which national rules could survive Articles 34 and 35,[70] and which could not.

This has been recognised more recently, with a string of decisions rephrasing the test and entrenching a requirement of disruption to 'market access.'[71] Discrimination was again dismissed as a factor[72] and the selling arrangement/product requirement dichotomy was apparently abandoned.[73] This resulted in a return to what seems to be a purely *effects*-based test akin to that in *Dassonville*,[74] which can adequately eliminate non-tariff barriers.[75] However, *Keck*'s spectre remains.[76] The fact that the Court has not expressly renounced the reasoning behind it leaves the avenue of another potential retreat open. This is not acceptable.

[67] Treaty on the Functioning of the European Union, arts 34, 35.
[68] See Barnard and Peers (n 16), 336, noting that the Member States are still *involved* in the making of rules which will apply to them, despite the practical loss of absolute self-governance.
[69] *Keck* (n 24).
[70] Treaty on the Functioning of the European Union, arts 34, 35.
[71] Case C-110/05 *Commission of the European Communities v Italian Republic (Trailers)* [2009] ECR 519.
[72] Case C-456/10 *Asociación Nacional de Expendedores de Tabaco y Timbre (ANETT) v Administración del Estado* [2012] ECLI:EU:C:2012:241.
[73] Case C-333/14 *Scotch Whisky Association and Others v Lord Advocate and Advocate General for Scotland* [2016] 2 CMLR 27.
[74] *Dassonville* (n 17).
[75] Laurence W. Gormley, 'Inconsistencies and Misconceptions in the Free Movement of Goods' [2015] 46 European Law Review 6, 925: 'market access' adds little to the original *Dassonville* test.
[76] *Keck* (n 24).

It is submitted that the law must stand clearly – integration is paramount to the EU and the concept of 'market access' must be emphasised and clarified.[77]

[77] Objective clarification is particularly important. See Snell, 'The Notion of Market Access: A Concept or a Slogan?' [2010] 47(2) CMLR 437, where the author argues that the term's lack of content may result in yet more unprincipled decisions based on 'intuition.'

The Key Shifts in EU-UK Trade After the Trade and Cooperation Agreement – an Artificial Separation Led by Europe

Eve Poyner

Introduction

The most conspicuous, and arguably detrimental, key change to EU-UK trading relations is that the UK is no longer part of the EU Single Market or Customs Union. *Prima facie*, the UK government fulfilled its plight to migrate to the lowest form of economic integration with the EU, but it is submitted that, beyond the looking glass, there is a different narrative entirely. This narrative, heavily shielded by Boris Johnson's assertion of 'restoring national sovereignty'[1] and the Conservative party's 'Global Britain'[2] agenda, hides the inextricable ties the UK will have to the EU, in terms of trade, for the foreseeable future. This is why the key shifts in relations have changed on paper but not materially. The global reach of EU law and overall dominance of the Union's trade policies have meant that, despite its departure, the UK will irrevocably be aligned to the EU without reaping the benefits of membership. Even future British endeavours appear to be heavily 'Europeanised.' This article will first seek to explain how the EU's global influence prevents any meaningful British impact in terms of global trade. It will then briefly explore how, even domestically, the UK remains tied to European acquis. Lastly, it will analyse examples of the UK's new Free Trade Agreements to prove the fallacy of the 'Global Britain' agenda and continued eminence of EU law, across the channel.

The EU's Global Power

The sheer international reach of EU regulation, through the de facto 'Brussels Effect',[3] means that the UK will be unable to achieve the regulatory freedom they desire. This will lead to British businesses being forced to continue trading on EU standards for them to stay afloat in the global market. The 'Brussels Effect' is the mechanism by which the EU, by exercising its internal goals of greater European integration within the Union and harmonised regulations across all Member States, forces businesses to adopt such regulations to gain access to the large single market.[4] In addition, the EU continues to pursue the most stringent regulations which encourages non-divisibility of a company's production across their entire market. It is likely to be costly if a business adopts a different set of standards for different markets. If they maintain standards which differ from that of the EU, then they surrender access to the world's

[1] UK-EU Trade and Cooperation Agreement (30 December 2020) <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:22021A0430(01)&from=EN> accessed 15 February 2021.
[2] Nick Witney, ''Global Britain': Still Waiting for The Big Reveal' (*UK in a Changing Europe*, 17 June 2020) <https://ukandeu.ac.uk/global-britain-still-waiting-for-the-big-reveal/> accessed 6 May 2021.
[3] Anu Bradford, *The Brussels Effect* (OUP 2020).
[4] ibid.

largest single market.[5] Thus, as about half of British exports are sold to the EU,[6] businesses on the isle will continue to rely on the internal market and continue to adhere to EU rules in the foreseeable future. Notwithstanding the size of the Single Market itself, recent association agreements with third countries direct those countries to largely align their regulations with that of Europe.[7] Thus, in order to trade with these states, the UK will be forced to operate based on EU rules. This aligns with the suggestion that the UK will become a 'rule-taker' from Europe. For example, several states and industries aspire to the EU's high, global standard in data protection;[8] under the General Data Protection Regulation (GDPR)[9] Argentina and Japan, for example, have largely modelled their privacy laws on the GDPR, and companies such as Apple and Uber use a GDPR-complaint policy for their operations worldwide.[10] Because the UK relies heavily on its services sector and internet economy, which involves excessive data transfer to the EU, the government has sought to prove that its protection is equivalent to that of the GDPR.[11] Although there may be some divergences, the UK's approach to data protection will be close to that of the EU's. As a third country, in terms of data protection, it has granted adequacy to the EU data standard.[12] The spill-over effect[13] EU regulation means that the global economy for data protection is on a level playing field, originating in Europe, thereby showing very little shift in EU-UK trading relations. In other words, the UK will continue to trade on the same European terms; access to the single market is too lucrative to decline and third countries are also trading on the same terms.

It is therefore natural to assume that the only way in which Brexit would equal a profitable outcome is if the UK could regulate in a way that would give them a competitive advantage over the EU.[14] Chalmers argues that the UK could gain this advantage by subsidising exporters, but this is prohibited by the World Trade Organisation (WTO) and the Trade and Cooperation Agreement (TCA).[15] The idea of 'taking back sovereignty'[16] is seriously thwarted here – the UK cannot escape sanctions of the WTO or EU and the obvious global reach of EU law effectively forces the UK to adhere to EU regulations, otherwise surrendering the UK's largest trading partner. However, the UK could regulate in a way which gives British businesses an 'edge,' e.g., by lowering standards,[17] but, not only is this prohibited under the TCA, it would

[5] European Union, 'Trade' < https://europa.eu/european-union/topics/trade_en> accessed 6 May 2021.
[6] House of Commons Library, 'Statistics on UK-EU trade' (*House of Commons Research Briefing*, 10 November 2020) <https://commonslibrary.parliament.uk/research-briefings/cbp-7851/> accessed 8 May 2021.
[7] EU-Georgia Association Agreement [2014] OJ L261/4.
[8] Elaine Fahey, 'The Cross-Channel Reach of EU Law in the UK Post-Brexit' in Juan Santos Vara, Rames A Wessel and Polly Polak (eds), *The Routledge Handbook on the International Dimension of Brexit* (Routledge 2020), 330.
[9] Council Regulation 2016/679 of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data [2016] OJ L119/1.
[10] Bradford (n 3).
[11] ibid.
[12] David Erdos, 'Data Protection After Brexit: A New Switzerland?' (*UK in a Changing Europe*, 22 January 2021) < https://ukandeu.ac.uk/data-protection-after-brexit-a-new-switzerland/> accessed 6 May 2021
[13] Fahey (n 7), 331.
[14] Damian Chalmers, 'British Sovereignty Run by Europe' (*UK in A Changing Europe*, 29 December 2020) < https://ukandeu.ac.uk/british-sovereignty-run-by-europe/> accessed 7 May 2020.
[15] ibid.
[16] Trade and Cooperation Agreement (n 1).
[17] Chalmers (n 14).

impose higher costs on companies with significant activity across the channel.[18] However, even if UK law would have a 'material impact on trade' it would only work if Britain could become the global hegemon in areas untouched by the EU. As the EU is setting the international high standard in a particular industry, and because many trading partners follow this standard, it is almost useless for the UK to try to emulate this. If the UK were to legislate higher standards than that of the EU, European companies may begin to withdraw from selling to the British market due to extra expense, notwithstanding the costs it would impose on smaller British businesses.[19] A third member third state is unlikely to drop their standards to forego access to a trading bloc of half a billion consumers, and where tariffs would be lower if they comply with the bloc's standards to trade with the UK. Uniform production makes this impractical and costly.[20] While there is potential for the UK to become a 'rule-maker' in areas where the EU hasn't delved into, such as Artificial Intelligence (AI) or financial services,[21] they will continue to rely on the single market as their largest trading partner which is not without the constraint of the TCA. Therefore, we see 'British sovereignty run by Europe'[22] such that the tangible changes, as a result of Brexit, have done very little intangibly.

The UK's Domestic Ties

As the UK was a Member State of the EU for over 40 years, British laws, businesses, and industries are aligned with the EU acquis. Consequently, all trade policies in the UK originated in the EU. Although the European Union Withdrawal Act 2018[23] was known as the 'Great Repeal Bill,'[24] this is not the case. On exit day, the 2018 Act simply incorporated a 'snapshot' of EU derived legislation into UK law. Thus, all EU legislation effective on 31 January 2020 continues to be directly effective.[25] As a result, all British businesses will continue to operate and trade on European legislation and, unless subsequently repealed by Parliament, will be able to argue that such regulations are effective in the UK courts. Moreover, although now limited,[26] the UK courts have discretion to consult retained EU case law. Furthering Chalmers point of 'British sovereignty run by Europe.'[27] This is clear evidence that EU-UK trading relations are unlikely to change. Fahey also notes that, because the UK is the only country to have withdrawn from the EU, its businesses and trade bodies have EU laws built into every fibre of their operation.[28] Moreover, the further retention of European acquis through the 2018 Withdrawal Act[29] clearly reveals that the EU and UK will be on similar trading terms for the foreseeable

[18] Kenneth A. Armstrong, 'Regulatory Alignment and Divergence after Brexit' [2018] 25(8) Journal of European Public Policy 1099.
[19] ibid.
[20] Bradford (n 3), 278.
[21] Fahey (n 8), 331.
[22] Chalmers (n 14).
[23] European Union (Withdrawal) Act 2018 (EUWA 2018).
[24] Rowena Mason, 'Theresa May's 'Great Repeal Bill': What's Going to Happen and When?' *The Guardian* (2 October 2016).
[25] Paul Craig, 'Constitutional Principle, the Rule of Law and Political Reality: The European Union (Withdrawal) Act 2018' [2019] 82 Modern Law Review 319.
[26] European Union (Withdrawal Agreement) Act 2020 (EUWA 2020).
[27] Chalmers (n 14).
[28] Fahey (n 8), 338.
[29] European Union (Withdrawal) Act 2018 (EUWA 2018).

future. The globalisation of standards adds to the unhelpfulness of the UK repealing such EU measures, as do provisions in the TCA which forbid any provisions that may affect competitiveness between the two.

As previously stated, any divergences made in the UK parliament which distort competition between the EU and UK can be met with dispute settlement mechanisms which are heavily crafted upon an EU landscape.[30] Although there is no obligation to consider either party's domestic law (the TCA is based upon public international law),[31] there is a strong contention that EU law will be relevant to arbitration processes. For instance, the TCA's state aid provisions may be considered using EU law, which has a more substantial case law, and the UK's own state aid regime also fundamentally resembles the EU's. [32] This hints at dispute settlement being dictated with a slight favour to EU law.[33] Overall, if the UK does try to diverge at the domestic level, it may be met with an EU-flavoured solution which illustrates the little change within EU-UK trading relations, namely due to a lack of British autonomy.

Future UK Free Trade Agreements (FTAs)

Considering previous assertions herein enumerated; based on its own trade agenda, the UK's future trade deals with third countries looks bleak. A closer reading of the politically lauded free trade agreements shows that EU law continues to dominate UK trade which furthers the argument that there are little changes to EU-UK trading relations after the TCA.

Lazowski's, seemingly compelling, argument that the UK-Georgia Strategic Partnership and Cooperation Agreement[34] (a roll-over agreement), which accounts for less than 0.1% of British trade[35], was prioritised to serve the 'taking back control' narrative of the Conservative government.[36] However, more precisely, the agreement with Georgia almost exactly mirrors the EU-Georgia Association Agreement.[37] For example, Article 15 is a replica of its European counterpart, which covers cooperation on migration and asylum etc.[38] While this may be an incidental similarity, it helpfully outlines the conjecture that the UK will continue to conduct trade agreements on the same terms as the EU; Thus, displaying the very little change in trading relations and the UK's continued allegiance to the Union. Perhaps the most notable caveat is Article 227[39] which provides for the upholding of labour and environmental law standards. As the UK possesses the same standard as the EU for both fields, we again see the UK's continued allegiance to European standards. The UK had provisionally agreed this deal with Georgia prior

[30] Mark Konstantinidis & Vasiliki Poula, 'From Brexit to Eternity: The Institutional Landscape under the EU-UK Trade and Cooperation Agreement.' (*European Law Blog*, 14 January 2021) <https://europeanlawblog.eu/2021/01/14/from-brexit-to-eternity-the-institutional-landscape-under-the-eu-uk-trade-and-cooperation-agreement/> accessed 7 May 2020.
[31] TCA (n 1).
[32] Konstantinidis & Poula (n 30).
[33] ibid.
[34] UK-Georgia Strategic Partnership and Cooperation Agreement (21 October 2019) WT/REG440/N/1 <https://docsonline.wto.org.> accessed 7 May 2020.
[35] Adam Lazowski, 'Copy-pasting or Negotiating? Post-Brexit Trade Agreements between the UK and non-EU countries' in The Routledge Handbook on the International Dimension of Brexit (Routledge 2020).
[36] ibid.
[37] UK-Georgia Strategic Partnership and Cooperation Agreement (21 October 2019) WT/REG440/N/1 <https://docsonline.wto.org.> accessed 7 May 2020.
[38] ibid, art 15.
[39] ibid, art 227.

to the TCA but, strengthened by the commitment to a level playing field in the TCA, there are practically no changes with regards to trade relations in this line of standards. This provision could have been rejected on the UK's side prior to ratifying the TCA, but the decision to include it is significant in securing the reality of the UK's adherence to European policy.

Another example of a roll-over agreement which invariably aligns the EU, and the UK is the UK-Switzerland trade agreement,[40] which only covers 3 sectors. Perhaps it is Switzerland's close ties with the EU that are responsible for this alignment, but it again provides evidence of little to no change in relations. As a result of Switzerland's Mutual Recognition Agreement with the EU,[41] which applies to 20 sectors, Switzerland is devoted to 'legislative equivalence'[42] with the Union. Accordingly, Switzerland will be unable to commit to recognition of UK practises if they are different from that of the EU. However, one difference between the EU's deal with Switzerland, in comparison with the UK-Swiss deal, is the scope of mutual recognition. Owing to this principle, the three sectors contained are covered by the UK's international agreements, so they could be rolled over.[43] The rest are based on complete equivalence rules between Switzerland and the EU.[44] However, there is a temporary unilateral recognition based on EU recognition for UK imports from 13 other sectors.[45] Although such roll-over agreements serve to allow for continuity while the UK defines its own trade policy, they helpfully indicate that it is unlikely that the UK will be able to diverge too far without breaching their obligations with third countries. This highlights the little change in EU-UK trading relations, such that the UK's standards will continually be on the same level as the EU, even outside of the Union.

Conclusion

In conclusion, the *tangible* consequences of leaving the European Union have done very little to alter the *actual* state of trading relations between the EU and the UK. A combination of factors is responsible for this: the UK's lengthy membership, and so inevitably difficult legal disentanglement; the EU's global high standard of regulation, and their sizable single market. Part 1 of this essay explained how theories such as the 'Brussels Effect' compels the UK to adhere closely to EU regulations in order to stay afloat in the global market and become a 'rule-taker'[46] without a voice in the creation of those regulations.[47] Any diversion will ultimately lead to poor consequences on the UK's part. Part 2 explored how the UK's domestic ties have made it nearly impossible for the UK to ever escape the reign of the EU; legislating its own

[40] UK-Switzerland Trade Agreement (11 February 2019) WT/REG 437/N/1 <https://docsonline.wto.org> accessed 7 May 2020.
[41] EU-Switzerland Mutual Recognition Agreement [2002] OJ L114/369.
[42] Panos Koutrakos, 'International Trade after Brexit: Rollover Agreements Concluded by the UK.' (*EU Relations Blog,* 19 June 2020) <https://eurelationslaw.com/blog/international-trade-after-brexit-rollover-agreements-concluded-by-the-uk> accessed 6 May 2021.
[43] Peter Ungphakorn, 'The United Kingdom-Switzerland Trade Agreements' (*EU Relations Law*, 28 January 2021) < https://eurelationslaw.com/blog/the-united-kingdom-switzerland-trade-agreements#more-855> accessed 7 May 2020.
[44] ibid.
[45] Federal Department of Economic Affairs, Education and Research, 'Recognition of Conformity Assessment,' 1 January 2021 (Switzerland).
[46] Fahey (n 8).
[47] Bradford (n 3).

trade policy which cannot contravene the TCA, or other trade agreements for that matter, forces maintenance of a level playing field. Finally, part 3 explained how the UK's new FTAs appear to be 'Europeanised' and are unlikely to prove superiority unless more favourable terms are negotiated. Ultimately, the key changes of the TCA have not produced equivalent shifts in trading relations. Although there is time for the UK to develop its own defined trade policy, a lack of developed bureaucracy and the reasons outlined above show that this will not happen for the foreseeable future. In this bleak, post-Brexit landscape, the inevitable conclusion is that the UK will continue to slavishly bow down to Europe's unconscious sovereignty, without any say in its mastery.

Getting PIP'd Off: A Short Analysis of the Delays in and the Inadequacies of Personal Independence Payments and its Contribution to the Hostile Environment Towards Disabled People

Daniel Holt

Introduction

The welfare state in England and Wales is failing disabled people and people with health conditions ('DPPHC' hereafter) and contributing to a wider hostile environment experienced by this community. Cuts to social services and welfare reforms led to 'grave or systematic' violations of the rights contained in United Nations ('UN') Convention on the Rights of People with Disabilities.[1] The UN special rapporteur on extreme poverty and human rights adds that 'the social safety net has been damaged by drastic cuts'[2] and that the net 'has been systematically and starkly eroded, particularly since 2010.'[3]

The current system is far removed from the welfare state implemented after the Second World War, which was accompanied by commitments to economic growth and full employment.[4] A long-lasting decline began in the 1970s with the abandonment of these pledges despite continuing growth in public expenditure, income, and wealth inequality.[5] The Great Recession and the Conservative and Liberal Democrat coalition government in 2010 followed. The newly established executive condemned the nation to austerity and the Welfare Reform Act 2012.[6] These reforms included the social and private housing sector, contributory and non-contributory benefits, tax credits and out-of-work and in-work benefits and have affected all segments of the population, including children, women, single parents, older persons, and disabled people.[7]

The reforms resulted in a drastic reduction in the public funds supporting DPPHC, such as benefits and budgets covering extra costs associated with disability. The government simultaneously adopted rhetoric that DPPHC are 'benefit scroungers'[8] who take advantage of the social security system and a populist discourse around welfare dependency and 'making work pay.' This has been reinforced by newspaper headlines and television programmes such

[1] Convention on the Rights of Persons with Disabilities, 'Inquiry concerning the United Kingdom of Great Britain and Northern Ireland carried out by the Committee under article 6 of the Optional Protocol to the Convention' (*United Nations*, CRPD/C/15/R.2/Rev.1, 6 October 2016) para 113.
[2] Philip Alston, 'Visit to the United Kingdom of Great Britain and Northern Ireland: Report of the Special Rapporteur on Extreme Poverty and Human Rights' (*United Nations,* A /HRC/41/39/Add.1, 23 April 2019). para 29.
[3] ibid.
[4] Christopher Pierson, *Beyond the Welfare State: Beyond the Welfare State* (3rd edn, Polity Press 2006).
[5] ibid.
[6] Jed Boardman, 'Dismantling the social safety net: social security reforms, disability and mental health conditions' [2020] BJPsych Bulletin 44, 208–212, p.209 <https://www.cambridge.org/core/journals/bjpsych-bulletin/article/dismantling-the-social-safety-net-social-security-reforms-disability-and-mental-health-conditions/B73ED8C6EBEBD1A663EE964146451705> accessed 24 January 2022.
[7] Convention on the Rights of Persons with Disabilities (n1) para 34.
[8] Frances Ryan, *Crippled: Austerity and the Demonization of Disabled People* (Verso 2019), 36.

as *Benefits Street*,[9] which portray disabled people as being dependent or making a living out of benefits, committing fraud as benefit claimants, being lazy, and putting a burden on taxpayers who are paying 'money for nothing.'[10] Public support is lacking as a result of the welfare state being perceived as a handout for people who have no desire to work.[11] DPPHC have also endured an increase in hate crimes and aggression.[12] Over 9,200 disability hate crime reports to police were reported between April 2020 and April 2021, with half classified as 'violent'-involving assault and possession of weapons.[13] The rate of repeat offending has drastically increased by 89% on the 2019/20 statistics,[14] which suggests perpetrators are not being deterred and disability hate crime is tolerated by society. The COVID-19 lockdowns, which were introduced to reduce its spread by heavily restricting when one can leave the home, offered no refuge as online abuse towards disabled people increased by 52%.[15]

This essay focuses on the contribution of Personal Independence Payments ('PIP') to the hostility towards DPPHC through insufficient funding, severe delays, and inadequacies in its assessments. PIP is a non-means-tested benefit that was introduced by the Welfare Reform Act 2012 to replace Disability Living Allowance ('DLA') and aimed at reducing £13.5bn DLA expenditure by 20%.[16] The government removed the lowest support rate, which was previously in the care component of DLA, and narrowed eligibility to supposedly appease its cost-cutting agenda. In addition, 48% of people previously on DLA were no longer supported by PIP or had their payments lowered.[17] This has left many disabled and unwell people without vital income.

PIP cannot be described as anything other than 'nothing short of a fiasco.'[18] Its financial aims have woefully failed, as the Office for Budget Responsibility estimated in March 2016 that expenditure only reached a reduction of 5%.[19] The regime survives, however, while DPPHC are deprived of basic support. This essay sheds light on the severity of the delays and hostility towards DPPHC through administrative injustice. PIP is inherently incapable of protecting the vulnerable; it is the embodiment of grave erosion of the rights, independence, and dignity of DPPHC. Claimants are often in a precarious position psychologically owing to the rigours of the constant assessment process.[20] One example is Susan Margaret Roberts, who

[9] ibid, 31.
[10] Convention on the Rights of Persons with Disabilities (n1) para 84.
[11] Ben Baumberg, 'The stigma of claiming benefits: a quantitative study' [2016] 45 Journal of Social Policy 181, 99.
[12] Convention on the Rights of Persons with Disabilities (n1) para 85.
[13] Press release section, 'Lockdown's trigger surge in disability hate crime' (Leonard Cheshire, 2021) <https://www.leonardcheshire.org/about-us/our-news/press-releases/lockdowns-trigger-surge-disability-hate-crime> accessed 24 January 2022.
[14] ibid.
[15] ibid.
[16] Michael Buchanan, 'Pensions stance creates budgetary mayhem' (BBC, 21 March 2016). <https://www.bbc.co.uk/news/uk-politics-35864438> accessed 24 January 2022.
[17] ibid.
[18] Rt Hon Margaret Hodge MP, Chair of the Committee of Public Accounts, stated upon the publication of House of Commons Committee of Public Accounts, *Personal Independence Payments* (HC 280 2013-14).
[19] Office of Budget Responsibility, *Economic and fiscal outlook* (Cm 9212, March 2016), para 4.116 and 1.9; The most recent figures indicate that £16.5bn will be spent in 2021-22, £18.3bn in 2022-23, £19.8bn in 2023-24, £21.4bn in 2024-25, and £23.4bn in 2025-26.
[20] See generally Frances Ryan (n8).

was discovered dead by a care worker at her warden-assisted flat surrounded by letters telling her that she would not be entitled to PIP.[21] I write this essay in their honour.

Chapter One: Delays in the Initial Application

PIP experienced significant delays from the outset because the process was poorly designed. The High Court in the 2015 case of *R (Ms C & Mr W) v Secretary of State for Work and Pensions* [SSWP][22] ruled that delays of thirteen and ten months in paying benefits to two DPPHC were 'unlawful.'[23] SSWP argued that sufficient steps had been taken by increasing the number of assessors, assessments, assessment centres and administrative staff.[24] Guidance, communication and forms were also improved.[25] The court was rightly unconvinced and provided twelve reasons for finding against the SSWP. These cover severe delays, the vulnerability and additional needs of the claimants, unnecessary demands in the process, assessment providers' lack of capacity, distress caused, the insufficiency of back payments, and the SSWP's unreasonable behaviour.[26]

The difficulty and hostility endured by DPPHC navigating PIP and the wider social security system achieved sufficient severity for Ian Duncan Smith, the contemporary SSWP, to resign from his role. He stated that these reforms were underfunded[27] and social security for disabled people had been cut too far when he provided reasons for his resignation.[28] The DWP relies on the excuse that it is financially restrained in its attempts to make improvements.[29] It argued that it had to proceed with the transition from DLA to PIP because backtracking would have been a departure from the Welfare Reform Act 2012, resulting in legal ramifications and impacting those that benefit from PIP.[30] However, it appears that this system was designed to limit the financial support available to DPPHC without sufficient understanding of the wellbeing of those who need to be, and have been, processed for the benefit.

The National Audit Office ('NAO') identified issues with the introduction of PIP and concluded that the DWP failed to leave time to assess how the length of the claims process would be affected by increased volumes.[31] The NAO also found that the lack of time did not allow for an evaluation of the delays in 'assessments before inviting new claims from across

[21] John Pring 'PIP investigation: Woman took her own life two days after learning of failed PIP appeal' (Disability News Service, 16 February 2017) < https://www.disabilitynewsservice.com/pip-investigation-woman-took-her-own-life-two-days-after-learning-of-failed-pip-appeal/> accessed 24 January 2022.
[22] *R (Ms C & Mr W) v Secretary of State for Work and Pensions* [2015] EWHC 1607.
[23] ibid., [93].
[24] ibid., [68].
[25] ibid.
[26] ibid., [94].
[27] Editorial team 'In full: Ian Duncan Smith resignation letter' (BBC, 18 March 2016) <https://www.bbc.co.uk/news/uk-politics-35848891?fbclid=IwAR0YjqGndqCemfpKwVSPzQ5N26X4ghHjMx2SAoFkbPQnrzJqMkuvgDWA7KA> accessed 24 January 2022.
[28] ibid.
[29] ibid.
[30] *R (Ms C & Mr W)* (n22), [75].
[31] Comptroller and Auditor General, 'Personal Independence Payment: Early Progress' (National Audit Office 27 February 2014) para 3.15 <https://www.nao.org.uk/wp-content/uploads/2014/02/Personal-independence-payment-early-progress.pdf> accessed 24 January 2022.

the country in June 2013.'[32] The DWP did not have sufficient time to 'identify problems before introducing natural reassessment of Disability Living Allowance claims in October 2013' either.[33] It had 'only two months to resolve problems before volumes would increase again to around 55,000 claims per month'[34]

The Public Accounts Committee findings were equally critical and highlighted the causes of 'significant delays to benefit decisions and a growing backlog of claims': [35]

> *'The [Department] […] did not pilot the benefit process [… and] significantly misjudged the number of face-to-face assessments […] and the time their assessments would take.'*[36]

The most recent official statistics[37] show that the length of the PIP claim process, from registration to a decision, is 112 days (16 weeks).[38] The process has been prolonged during the COVID-19 pandemic to 38 days more than the 74-day original timeframe for a new claim.[39] 112 days is the revised target figure given in evidence by the Minister for Disabled People to the Work and Pensions Committee (WPC).[40] The WPC described this figure as unacceptable as applicants should not be waiting six months or more for a decision that can enable financial support towards the additional costs incurred by DPPHC.[41] The WPC gave three recommendations for addressing the delays in the process. Firstly, DWP should 'closely examine its own systems and that it work[s] with the contracted providers to resolve the current dire situation' and 'penalty clauses should be invoked where necessary.'[42] Secondly, DWP needs to clear the claims backlog. Thirdly, DWP should reduce the length of the process for new claims to the expected 74 days, before it reassesses existing DLA claims.[43] The DWP has failed to meet these recommendations as the expected 74 days continues to be unmet. The pandemic fails to provide justification as severe delays existed before the global public health crisis. The recorded length for March 2020, when the DWP introduced the measures responding to COVID-19, was 147 days;[44] double the 74-day timeframe[45] and 35 days longer than the WPC figure of 112 days.

[32] ibid.
[33] ibid.
[34] ibid.
[35] Committee of Public Accounts, *Personal Independence Payments* (HC 280 2013-14), 9.
[36] ibid.
[37] Department of Work and Pensions, Personal Independence Payment: Official Statistics to October 2021' (*Department for Work and Pensions* 14 December 2021).
[38] Under normal rules, not claims not being processed under 'special rules for terminal illness (SRTI).
[39] Comptroller (n34) para 1.4.
[40] *R (Ms C & Mr W)* (n22) [37].
[41] Work and Pensions Committee, *Monitoring the performance of the Department for Work and Pensions in 2012-13* (HC 2013–14, 1153), para 48.
[42] ibid.
[43] ibid.
[44] Department for Work and Pensions, *Personal Independence Payment: Official Statistics (data to January 2020, March 2020)*, 4.
[45] Comptroller (n31) para 1.4.

The government admits that PIP exists to help meet the additional costs of being a DPPHC.[46] Excessive delays in awarding PIP place claimants who are managing health conditions into further precarious emotional and financial situations as with wrongly denied payments. The government has failed to rectify the delays that were apparent from the introduction of PIP, and it has no plans to address the process length as the WPC recommended. The delays in the process are absent from both the National Disability Strategy[47] and Shaping Future Support: The Health and Disability Green Paper.[48] This indicates that the government has no desire to act. In a hostile environment where adverse and derogatory rhetoric is commonplace, the anguish experienced by members of a community that already face unemployment, exclusion and discrimination furthers the cost cutting agenda. The status quo continues despite repeated criticism, whilst many applicants are only able to access financial support to which they are entitled by enduring the request for Mandatory Reconsideration (MR) and possibly an appeal to the Social Security and Child Support Tribunal (SSCS Tribunal).

Chapter Two: Assessments

Applicants are often invited to an assessment in the initial process as part of determining how their condition affects them and their eligibility for PIP. It stands to reason that the assessment process is a key component of the hostile environment towards DPPHC. The antagonism begins with the application of an overtly narrow eligibility criteria that fails to consider conditions in their totality and different levels of pain, difficulty, or distress.[49] The PIP criteria fail to recognise or represent the challenges faced by DPPHC. The criteria are seemingly designed to limit eligibility and appear to be part of an environment that perceives this community as 'scroungers.' Assessors assume that people are lying about or exaggerating their conditions,[50] and many applicants are considered fraudulent.[51] Assessors often fail to appreciate the challenges of managing impairments and/or health conditions and applicants' specific needs.[52] Applicants are treated as 'wholly inferior' and second-class citizens as assessors have been known to ask questions they cannot understand and prevent them from using the toilet.[53]

The assessment's approach to eligibility makes DPPHC feel chastised. The fact that assessors appear unable to adequately assess the relevance of applicants' conditions, due to a lack of extensive medical qualifications, makes it seem as if the assessments are designed to be uncomfortable.[54] The rigid questioning caused by a lack of understanding means the assessment fails to capture an accurate picture of applicants' conditions, especially since

[46] Department for Work and Pensions, *PIP Assessment Guide (*24 January 2022), para 1.1.1.
[47] Department for Work and Pensions, *National Disability Strategy* (CP 512, 2021).
[48] Department for Work and Pensions, *Shaping Future Support* (CP 470, 2021).
[49] Jen Durrant, 'Access Denied: Barriers to Justice in the disability benefits system' (Zacchaeus 2000 Trust, 2018) <https://www.z2k.org/wp-content/uploads/2018/09/Z2K_disability_report_2018_Final_singlesheet-1.pdf> accessed 24 January 2022, 14.
[50] Ben Baumberg Geiger, 'A Better WCA is Possible' (Demos, 2018) <https://demos.co.uk/project/a-better-work-capability-assessment/> accessed 24 January 2022.
[51] Jen Durrant (n49), 38.
[52] ibid.
[53] ibid.
[54] ibid.

assessors do not know when to ask follow-up questions.[55] This is furthered by refusing to sufficiently recognise non-specialist evidence as 'valid', which results in many claims being refused.[56] This discriminates against the DPPHC who cannot access specialist treatment due to a lack of available resources in health and social care. For example, evidence from social services, support agencies, friends, neighbours, and carers are dismissed as sub-standard despite being equally important.[57] The end result is a debasing process that makes applying for financial support difficult and uncomfortable.

Chapter Three: Mandatory Reconsiderations

The MR process was introduced to prevent flawed assessment decisions from going to the SSCS Tribunal. The PIP process is extended by a median time of 78 days for these applicants.[58] The waiting time is 30 days longer than the excessive 48.5 day waiting times recorded in March 2020 as coronavirus was taking effect.[59] The DWP argues that delays are a result of steps taken to '"enhance" the system [by gathering] further evidence from claimants and make more accurate decisions sooner.'[60] This does not reflect the experience of PIP as the process has been rightly criticised for failing to give enough support, guidance and assessment on eligibility.[61]

Regardless of its purpose, the MR process deters many applicants from pursuing their claim and obtaining much needed support. Applicants ended their pursuit of PIP following a negative decision in 23% of cases, which is disheartening since 21% of MR requests are successful.[62] Therefore, a significant number of these applicants will be wrongly denied the benefit. Making a MR request involves a significant administrative burden as applicants are only given one month from receiving their assessment decision to make their request. The difficulty in collating, preparing, and submitting a MR request without sufficient support intimidates and deters applicants, especially since they are managing conditions and navigating daily life.[63] Applicants often struggle to obtain evidence[64] and understand the process[65] and describe assessments as 'shockingly poor and dishonest.'[66]

The MR process was introduced to address the inaccuracies in the initial decision-making procedure and to provide an opportunity for applicants to restate their case. However, providing further steps to a broken system only prolongs the hardships experienced by DPPHC. The wait for vital financial support is lengthened and the budget strains are heightened. The

[55] ibid.
[56] ibid.
[57] ibid.
[58] ibid.
[59] Department for Work and Pensions (n44), 8.
[60] Dan Bloom, 'DWP disabled benefit appeal times skyrocket as PIP wait doubles to all-time high' (The Mirror, 2019) < https://www.mirror.co.uk/news/politics/dwp-disabled-benefit-appeal-times-20046721> accessed 24 January 2022.
[61] Department for Work and Pensions, *Early process evaluation of new claims for Personal Independence Payment* (Research Report No 867, July 2014), 3.
[62] Department for Work and Pensions (n 37).
[63] Jen Durrant (n 49), 44.
[64] ibid.
[65] ibid.
[66] John Pring, 'PIP investigation: Council probes 'shockingly poor and dishonest assessments' (The Disability News Service, 16 March 2017) < https://www.disabilitynewsservice.com/pip-investigation-council-probes-shockingly-poor-and-dishonest-assessments/> accessed 24 January 2022.

burdens of proving one's entitlements are continued, which adds to the daily difficulties experienced by DPPHC. It seems as if adding an ineffective MR process illustrates the unwillingness to address the long-term concerns with PIP as it prioritises its cost-cutting objectives over the rights of DPPHC.

Chapter Four: Appeals to the Social Security and Child Support Tribunal

The applicant only has the option of lodging an appeal with the SSCS Tribunal if the MR request is unsuccessful. This option is undertaken by 32% of applicants[67] and presents the first opportunity for the application to be independently assessed outside of the DWP. Lodging an appeal is off-putting and adds to the hostile environment. Applicants are faced with navigating a system they do not understand without the specialist skills required to effectively argue a case through oral and written submissions. Completing the required steps to lodge and participate in an appeal is strenuous and imposes physical, mental, and emotional strains, which should not be experienced by those seeking necessary financial support. Lawyers estimate it can take them up to 25 hours' work to submit an appeal despite having these skills and understanding.[68] Most applicants will not be able to access legal aid to pay for a lawyer to help formulate their case due to the £350m cuts.[69] Legal aid awards for social security cases fell 99% between 2012 and 2017.[70] The vast number of people who need free legal advice results in advice charities and pro bono lawyers often having limited capacity to help. Consequently, thousands of people who are incorrectly refused PIP must battle in the SSCS Tribunal without any financial support for legal advice or representation at all. The task can be difficult and overwhelming for applicants.[71] Lodging an appeal extends the emotional and financial hardships presented by the PIP process. PIP applicants wait a further 182 days[72] for the tribunal to arrive at a conclusion that should have been reached at the end of the initial process. There is also an additional 4 to 6 weeks' wait to receive payments.[73]

The strain of lodging an appeal forms part of the hostile environment experienced by DPPHC. The fact that 73% of applicants who lodge an appeal with the SSCS Tribunal are successful is damning.[74] Delayed support results in increased emotional anguish, time pressures and financial instability. This is accompanied by stress, anxiety, and distress caused by the unduly challenging, lengthy, and humiliating process. The success rate would be more condemning if more applicants were able to appeal. Barry et al concluded that of those not lodging an appeal, 37% said the main reason was that the process would be too stressful, 20% did not expect the award to change, and 20% were too unwell.[75] These applicants will most

[67] Department of Work and Pensions (n 37).
[68] Editorial team 'Legal aid: UK's top judge says cuts caused 'serious difficulty' (BBC, 27 December 2019) <https://www.bbc.co.uk/news/uk-50923289> accessed 24 January 2022.
[69] ibid.
[70] Jen Durrant (n 49), 30.
[71] ibid.
[72] ibid.
[73] Citizens Advice, *Challenging a PIP decision - the tribunal hearing,* (Citizens Advice) <https://www.citizensadvice.org.uk/benefits/sick-or-disabled-people-andcarers/pip/appeals/your-hearing/> accessed 24 January 2022.
[74] Department of Work and Pensions (n 37).
[75] Barry et al, *Personal Independence Payment Claimant Research—Final Report* (Department of Work and Pensions Research Report 963, September 2018), 18.

likely go without due financial support, and forgo the necessary perseverance required for the cruel and lengthy process. The result is further hardship amongst an already impoverished community.

Chapter Five: The Poverty Impact

The UK Government has claimed that PIP is designed to mitigate the extra costs faced by individuals with the greatest barriers to participation.[76] The DWP has also stated that the welfare system provides financial support and is revised to ensure the system remains fit for purpose[77] through engagement with disabled people.[78] On the other hand, the experience of those who engage with the PIP process is vastly different. PIP is a cost-cutting exercise that is failing to address the additional costs of living for disabled people, which is estimated to be £583 per month more than the provisions of PIP and other disability-related social security.[79] One in five disabled adults have outlays of over £1,000 per month as they endure unavoidable expenditure to cover specialist home or vehicle adaptations, therapies, equipment, sensory equipment or adapted toys for children, increased utility bills, taxis because of inaccessible public transport and higher insurance costs.[80] The pandemic has increased the financial burdens on disabled people. A significant 31% of disabled people reported difficulties accessing groceries, medication, and essentials during the pandemic in September 2020 compared with 12% of non-disabled people.[81] Equally, 36.7% of disabled people reported a rise in household outgoings; only 22.8% of non-disabled people documented such difficulties.[82]

PIP is not providing the financial support it is designed to deliver and the result is a state of anguish for many DPPHC. The United Nations Committee on the Rights of Disabled Persons ('UN Committee') nevertheless found a violation of Article 28, which requires 'an adequate standard of living and to social protection, […] poverty reduction and social protection programmes and financial assistance.' [83]

Evidence shows there is a 'new destitution' in that those who would have avoided absolute destitution with support from social security are now unprotected.[84] A 2018 Joseph Rowntree Foundation ('JRF') study found that the lack of financial support means that 1.5

[76] Department for Work and Pensions, *The UK's 2019 response to select concluding observations of the UN Committee on the Rights of Persons with Disabilities* (October 2019), para 12.
[77] ibid., para 8.
[78] ibid.
[79] Editorial team, 'Disability price tag report' (Scope, 2019) <https://www.scope.org.uk/campaigns/extra-costs/disability-price-tag/> accessed 24 January 2022.
[80] ibid.
[81] Catherine Putz and David Ainslie, 'Coronavirus and the social impact on disabled people in Great Britain' (Office for National Statistics, 11 November 2020) <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/disability/articles/coronavirusandthesocialimpactsondisabledpeopleingreatbritain/september2020> accessed 24 January 2022.
[82] ibid.
[83] Neville Harris, 'The rights of persons with disabilities: time for a review of benefits?' [2019] Journal of Social Security Law 26(3), 101-104, 101; see also Convention on the Rights of Persons with Disabilities (n 1) paras 58-59.
[84] Patrick Butler, 'Welfare spending for UK's poorest shrinks by £37bn' (The Guardian, 23 September 2018) <https://www.theguardian.com/politics/2018/sep/23/welfare-spending-uk-poorest-austerity-frank-field> accessed 24 January 2022.

million citizens have insufficient income to buy essentials or meet their core material needs.[85] DPPHC are amongst the most affected with 4 million disabled adults living below the breadline.[86] Thus disabled people represent one third of all adults in poverty.[87] Worse still, 1 in 5 working-age disabled people are living in 'severe material deprivation', which means that the individual is unable to afford four of the following: rent/mortgage/bills, adequate heating, unexpected expenses, meat/protein, a holiday, a television, a washing machine, a car and a telephone.[88] Specifically, malnutrition is a significant concern as 1 in 5 disabled people have to skip meals and lack essential nutrients.[89] Similarly, the inability to afford adequate heating results in 1 in 6 disabled people wearing a coat indoors.[90]

The PIP process makes a significant contribution to the deprivation by failing to provide sufficient financial support and creating an adverse system whereby DPPHC encounters unnecessary struggle and hardship. This was aptly described by the Public Accounts Committee:

> *'The delays […] are causing unacceptable [...] stress. […C]laimants [...] resort to loans, food banks and discretionary housing payments.'*[91]

Conclusion

PIP is failing to cover the additional costs incurred through being a DPPHC and this community is at a tangible risk of poverty as a result. This is only one way in which PIP contributes to the hostile environment towards this community. Accessing essential financial support is delayed significantly by a system that has failed to deliver within the 74-day target from the outset. The DWP has received sufficient advice and recommendations from the judiciary and parliamentary bodies since its inception that would reduce the processing times. The DWP opted to introduce a MR process instead, which only adds to the lengthy proceedings, to prevent cases from appearing before the SSCS Tribunal. The high success rate of appeals suggests that the process is largely ineffective in addressing flawed decision-making. This causes unnecessary stress and financial hardship for those navigating PIP. The fact that DWP has no intention of improving PIP is significant in demonstrating how DPPHC are treated within the system.

The assessments taking place during the initial process add to the harsh experience and the wider adversity experienced by DPPHC. The starting point seems to be that applicants are assumed to be ineligible and are treated as 'second-class citizens' who are 'scrounging' from the state. This results in assessors assuming that people are lying about or exaggerating their

[85] Patrick Butler, Mathew Taylor and James Ball, 'Welfare cuts will cost disabled people £28bn over five years' (The Guardian, 27 March 2013) <https://www.theguardian.com/society/2013/mar/27/welfare-cuts-disabled-people> accessed 24 January 2022.
[86] ibid.
[87] ibid.
[88] Patrick Butler, 'Destitution is back: And we just can't ignore it' (The Guardian, 3 July 2016) <https://www.theguardian.com/society/2017/jul/03/desititution-ignore-welfare-cuts> accessed 24 January 2022.
[89] ibid.
[90] ibid.
[91] Public Accounts Committee, *Personal Independence Payment* (HC 2013–14, 280), para 16.

conditions. The assessment forms part of the hostile environment endured as a result. Applicants are left feeling dehumanised and often guilty for seeking financial support.

The sad reality is that PIP and the wider social security system is contributing to the hostile treatment of DPPHC. Public support is dwindling and hate crime is increasing. This is reinforced by the government and sections of the media in adopting a rhetoric that people who do not work are 'benefit scroungers' who want 'money for nothing.' The system is underfunded and inadequately designed to treat applicants respectfully and provide support when it is needed. DPPHC are a valued part of our society and provide skills, narrative and existence that enrich the UK. This community deserves a social security system that protects them from poverty. Instead of 'making work pay', let's make 'life worth living.'

Principles of Party Autonomy and Limited Judicial Intervention in Contrast: Does the English Arbitration Act Strike a Fair Balance?

Zulker Nayeen

Abstract

In arbitral jurisprudence, a sense of competition between the principles of party autonomy and limited intervention by the court always attracts the concentration of different commentators. Previously, the English arbitration practises were immensely criticised for its interventionist attitude into the arbitration proceedings and arbitral awards. Limiting the judicial intervention, as one of the factors, prompted the Government to enact the new Arbitration Act in 1996. This Act enshrines both the competing issues as general principles of arbitration. However, does the Act establish the esteemed balance? A plain reading of the concerned provisions of this Act reveals that there are two restrictions on party autonomy: first, compliance to some mandatory provisions during undertaking an arbitration agreement; and second, necessary safeguards for upholding public interest. Does the Act allow the court to exercise its authority beyond these two restrictions? If yes, then how does this Act strike a fair balance between the principles of party autonomy and limited judicial intervention? This study analyses these issues.

Introduction

In legal parlance, party autonomy is a recognised principle of arbitration which demands proper respect from national courts. Prior to the enactment of the Arbitration Act 1996,[1] the English arbitration practises were immensely criticised for providing a wide scope for court intervention in arbitral proceedings. Consequently, the Act was enacted giving English arbitration law an entirely new face, a new policy, and a new foundation. It aims to replace the earlier broader scope of judicial intervention and embody a new balance of relationships between the parties, advocates, arbitrators, and the court.[2] The principles of party autonomy and limited judicial intervention are the offshoots of this balance. However, does the Act strike a fair balance between these two general principles? Before answering this question, it is essential to know which elements constitute such a fair balance. The principle of party autonomy in any arbitration agreement aims for an independent and private arrangement of dispute resolution among the parties.[3] It needs support from the judiciary for its proper enforcement. The judiciary supports the parties' autonomous arrangement of dispute resolution if it aligns with the basic judicial principles and public interest of the country. When any such autonomous arbitral arrangement complies with the principles and public interest issues of the

[1] Hereinafter referred to as the Act.
[2] Lord Mustill and Stewart C Boyd, *Commercial Arbitration: 2001 Companion Volume to the Second Edition* (Butterworths 2001) Preface para 1.
[3] Yas Banifatemi, 'Chapter 19: The Law Applicable in Investment Treaty Arbitration' in Katia Yannaca-Small (ed), *Arbitration Under International Treaty Arbitration* (2nd edn, OUP 2018) 485.

country, and the judiciary restrains itself from intervening into the matter, it can be called that the arbitral system and judiciary are maintaining a fair balance. This study aims to find out whether a fair balance is being maintained between the principle of party autonomy and the principle of limited judicial intervention under the present Act. To do so, two aspects were explored in this research: the extent of party autonomy which the Act allows; and ways in which the Act refrain the court from intervening into party autonomy.

The Extent of Party Autonomy of Party Autonomy Under the Arbitration Act 1996

The doctrine of party autonomy, which was first developed by academics, has gained extensive acceptance in national legal systems.[4] It seems necessary to identify what party autonomy is and how it comes into existence, before engaging in further discussion. Generally, party autonomy is the discretionary power of the parties whereby they can agree upon the laws and procedures to be applied in resolving their dispute arising out of any agreement.[5] An arbitration agreement derives its power from party autonomy.[6] Arbitral tribunal owes its existence to the agreement of the parties and, in applying the law chosen by the parties, an arbitral tribunal is simply carrying out their agreement.[7] Therefore, party autonomy comes into sensible existence through an arbitration agreement.

For determining the extent of party autonomy, it is necessary to understand the limit of discretion the parties may incorporate in the arbitration agreement. Lord Mustill and Boyd in this regard stated, '[p]arty autonomy gives the parties and their lawyers the opportunity to control all aspects of the proceedings, however unsuited to the nature of dispute, and however wasteful in terms of money, time and effort the agreed method might be...'[8] Therefore, the essence of party autonomy includes the parties' freedom to determine their governing law,[9] and prefer arbitrators and arbitral mechanisms with a view to resolving their dispute. Such freedom is expected to be 'respected in every way possible.'[10] This freedom and the concerned national law's respect to it reveal the extent of party autonomy under this particular national law. In this part, freedom of the parties in exercising autonomy under the Act will be discussed through analysing the concerned provision of the Act.

Section 1(b) of the Act sets 'party autonomy' as one of the basic principles of arbitration. It states that, 'the parties should be free to agree how their disputes are resolved, subject only to such safeguards as are necessary in the public interest.'[11] The reference to certain terms in this provision discloses that freedom of parties is subject to certain limitations. The first part of this provision, for instance, 'should be free to agree' signifies the parties' freedom to design

[4] Nigel Blackaby and others, *Redfern and Hunter on International Arbitration* (6th edn, Oxford University Press 2015) para 3.97.
[5] Banifatemi (n 3).
[6] Sunday A Fagbemi, '*The Doctrine of Party Autonomy in International Commercial Arbitration: Myth and Reality'* [2015] 6 (1) AFE Babalola University: J. of Sust. Dev. Law & Policy 222, 226.
[7] Blackaby (n 4) para 3.99.
[8] Mustill and Boyd (n 2) 26.
[9] Julian D M Lew, Loukas A Mistelis, and Stefan M Kroll, *Comparative International Commercial Arbitration,* (Kluwer Law International 2003) para 17-10.
[10] Elizabeth Shackelford, 'Party Autonomy and Regional Harmonisation of Rules in International Commercial Arbitration' [2005-2006] 67 University of Pittsburgh Law Review 897, 903.
[11] Arbitration Act 1996 (AA 1996), s 1(b).

the laws and proceedings of an arbitration while recording an agreement. This design is not unqualified because the words 'should be free' indicate an implied restriction in exercising wide autonomy because the word 'should' be a conditional one. It allows the exercise of discretion in agreeing to any matter, but such discretion is under the supervision of the judiciary. If the drafters of this legislation had used 'shall be free' or 'are free',[12] it would have indicated exclusive freedom to exercise party autonomy. Ideally, this provision does not allow exclusive party autonomy, rather it impliedly demands observance of some standards while drafting an arbitration agreement. Moreover, the second part of the provision stipulates an explicit restriction in the case of designing the arbitration arrangements and deciding the dispute. It legislates that party autonomy will be exercised subject to the principle of public interest. However, it is to be noted that the excuse of 'public interest' though explicit in the provision is not under exclusive scrutiny. Why does the principle of 'public interest' not seem to be an exclusive forbiddance? This is because there are some subsequent qualifying phrases in this section, such as 'such safeguards as are necessary.'[13] These conditional phrases mean that any agreed matter in an arbitration proceeding may seem to be against the public interest of any country including the UK. Despite this fact, the judiciary may consider that certain matters are not 'necessary' in public interest. It allows the court to have a proportionality test in determining any issue to be against public interest. This analysis reveals that parties have the autonomy to design the method of their dispute resolution, but this autonomy should be under implied or express scrutiny of the judiciary. If any matter of the arbitration proceedings is alleged to be against public interest, the court will not declare that matter *ultra-virus* right away. It will rather assess if the matter really is necessary to be declared to be against public interest. If it is found unnecessary to be declared to be against public interest, the court will not interfere in the process. It appears that the Act arranges a supportive attitude to the arbitration proceeding, rather than allowing exclusive judicial power to interfere in the proceeding.

Part I of the Act, which establishes party autonomy, begins with the heading 'arbitration pursuant to an arbitration agreement.' It stipulates that there must be an agreement to arbitrate wherein the parties will agree with their preferred arbitration proceeding. For being effective under English law, this agreement must be in writing or somehow recorded[14] rendering an intention to submit to arbitration any present or future dispute.[15] A well drafted arbitration agreement can exclude the jurisdiction of the courts and reflect the real needs to express the desire of the parties.[16] Part I of the Act further provides specific guidelines to be followed during recording or drafting an arbitration agreement. Parties cannot insert their desires on a whim and their autonomy is not unfettered while preparing the agreement too. This is because there are some provisions in the Act which require strict compliance in drafting any agreement. For example, certain sections of Part I of the Act are 'mandatory', in the sense that they cannot be overridden by agreement of the parties. The mandatory provisions of this Part are listed in

[12] As prescribed in the UNCITRAL Model Law on International Commercial Arbitration 1985, art 19(1).
[13] AA 1996, s 1(b).
[14] ibid., s 5.
[15] ibid., s 6.
[16] Ar. Gör. Şeyda Dursun, 'A Critical Analysis of the Role of Party Autonomy in International Commercial Arbitration and An Assessment of Its Role and Extent' [2012] 1 Yalova Üniversitesi Hukuk Fakültesi Dergisi 161, 168.

Schedule 1 of the Act and those are effective notwithstanding any agreement to the contrary.[17] Any agreement in contravention of the mandatory provisions will not be effective on the plea of party autonomy. The scope of party autonomy is thereby restricted through the mandatory provisions.[18] In case of any non-compliance to these mandatory provisions, the court has the full authority to declare any agreement contrary to these provisions null and void and having no effect.

Conversely, there are some 'non-mandatory' provisions which do not require strict compliance. Parties to any agreement have the discretion to avoid compliance to these non-mandatory requirements.[19] The majority of sections in Part I are in the 'non-mandatory' category[20] which eventually allow the parties to agree on their own arrangements.[21] Therefore, the parties have freedom to exercise autonomy on any matter that fall within the purview of non-mandatory provisions through an agreement regardless of its noncompliance with those provisions. The presence of non-mandatory provisions allows the parties ample discretion to agree on various issues of their arbitration arrangements. For example, the parties are allowed to agree in all matters relating to the constitution of a tribunal, powers of the tribunal, and powers of the court to arbitral proceedings. Moreover, the parties can agree on the time of commencing arbitral proceedings,[22] and fixing all procedural and evidential matters.[23] They also have the autonomy to constitute the arbitral tribunal,[24] and specify the functions of the arbitrators, chairman,[25] or the umpire.[26] The parties are even free to agree on the powers of the arbitral tribunal, such as deciding on its own substantive jurisdiction,[27] appointing experts, legal advisors or assessors,[28] making provisional awards,[29] and providing appropriate remedies,[30] interest,[31] and costs.[32] Parties also have the autonomy to specify the role of courts in some matters of arbitral proceedings, for example, they may agree on the extent of court's power in determining any preliminary point of law,[33] its powers to extending time,[34] and even excluding the court from entertaining appeal.[35]

The above discussion reveals the extent of autonomy which the parties may have in drafting an arbitration agreement under the Act. It appears that this Act allows the parties to frame all the relevant matters according to their needs, which include the commencement of

[17] AA 1996, s 4(1).
[18] Bruce Harris, Rowan Panterose, and Jonathan Tecks, *The Arbitration Act 1996: A Commentary* (5th edn, Wiley Blackwell 2014).
[19] ibid., para 4B.
[20] ibid., para 4C.
[21] AA 1996, s 4(2).
[22] ibid., s 14(1).
[23] ibid., s 34(1).
[24] ibid., ss 15(1), 16(1), 17(1), and 18(1).
[25] ibid., s 20(1).
[26] ibid., s 21(1), 22(1).
[27] ibid., s 30(1).
[28] ibid., s 37(1).
[29] ibid., s 39.
[30] ibid., s 48.
[31] ibid., s 49(1).
[32] ibid., ss 61, 62, 63, and 65.
[33] ibid., s 45.
[34] ibid., s 50.
[35] ibid., s 69.

arbitration, composition, powers, and functions of the arbitral tribunal and all procedural and evidential matters. Additionally, Part I of the Act legislates several default measures under the auspices of certain legal terminologies, for example, 'unless the parties otherwise agree', or 'unless otherwise agreed by the parties', or 'if or to the extent there is no such agreement.' The default provisions have been carefully drafted to provide a balanced and functional set of rules for nearly all the non-mandatory issues that might arise, which will be adopted simply through staying silent about the matter.[36] Therefore, agreeing to arbitration under the English law without specifying an intention contrary to the default provisions is one sort of party autonomy. The parties, in this case, are agreeing to the default procedure prescribed in the Act, similar to expressly exercising the discretion of party autonomy. It means that the parties are submitting them under the authority of the Act in case of any dispute settled through arbitration. It would mean that they are agreeing to the proceedings prescribed in the Act by default instead of choosing their own proceedings. Thus, the non-mandatory provisions seem to be making the notion of party autonomy unrestricted. It means the parties will not confront any restriction in case of exercising their autonomy of getting the dispute resolved through their chosen ways. Therefore, it reveals that the Act allows wide party autonomy, but this is subject to two major restrictions: i) principle of public interest and ii) mandatory provisions. Does the Act truly confine the court within these two restrictions or can courts take stances beyond these two? A discussion on this hypothesis would answer better whether the balance is fair enough. Since the parties know there are only two restrictions, limiting judicial intervention within the purview of these two can be considered a fair practice.

Confining the Judicial Intervention Within the Apparent Restrictions on Party Autonomy

The principle of limited court intervention in section 1(c) of the Act states, 'in matters governed by this Part the court should not intervene except as provided by this Part.' It is clear recognition of party autonomy and the desire to limit the court's role in arbitration so as to give effect to it.[37] The underlying philosophy is, where parties have agreed that their dispute should be resolved through arbitration, the court should not intervene except and to the extent necessary.[38] It is believed that most arbitration seated in England and Wales are conducted throughout, without any need for court involvement.[39] As was described by Aikens J in *Elektrim SA v Vivendi Universal SA & Ors* whereby the approach embodied in the Act is 'to give as much power as possible to the parties and the arbitrators, and to reduce the role of the courts to that of a supporter of the arbitration process up to an award being made.'[40] Despite having this supportive approach, in the following paragraphs, endeavours would be carried out to check whether the court confines itself to the principle of public interest and mandatory provisions

[36] Susan Blake, Julie Browne, and Stuart Sime, *A Practical Approach to Alternative Dispute Resolution* (3rd edn, OUP 2014).
[37] David St John Sutton, Judith Gill, and Matthew Gearing, *Russell on Arbitration* (24th edn, Sweet & Maxwell 2015).
[38] Blake, Browne, and Sime (n 36), 26-61.
[39] Kieron O'Callaghan and Jerome Finnis, 'Chapter 20: Support and Supervision by the Courts' in Julian D. M. Lew and Harris Bor (eds), *Arbitration in England, with chapters on Scotland and Ireland*, (Kluwer Law International 2013), 20-21.
[40] [2007] EWHC 571 (Comm), [71].

through discussing the court's role in assisting arbitral proceedings and rectifying serious injustice.

Assistance to Arbitral Proceedings

It is said that the English court's approach has considerably shifted to support the notion of party autonomy. The courts in general have either adopted a broad and more flexible approach or have applied the Act strictly to reflect parties' intentions.[41] In the earlier part, mandatory provisions have been found as the apparent restrictions on party autonomy. However, it is not true because these provisions have been designed to support arbitral proceedings. The following points would exemplify the mandatory provisions' support towards party autonomy.

1. An arbitration agreement is the document which reflects the parties' desire to arbitrate any 'present or future dispute.'[42] But it might happen that 'a party may bring court proceedings in breach of an existing arbitration agreement.'[43] To prevent a party from breaching such agreement through bringing court proceedings, section 9(1) of the Act allows the other side to apply for a stay of those court proceedings[44] unless they are content to forego their right to have the dispute referred for arbitration and instead choose to defend the action before the court.[45] Where it is feared that proceedings are about to be commenced in any foreign country, it may be possible to apply for an anti-suit injunction.[46]

2. Party autonomy authorises the parties to agree on the prospective timeframes for having a dispute settled through arbitration. For example, the parties may include in their agreement within how many days of arising a dispute they would commence the arbitration,[47] how many days the tribunal may take to pass an award,[48] and how many days the tribunal could take to complete the arbitral proceedings.[49] However, the mandatory provisions authorise the court to intervene in the matters of time-frame at the instance of any party if the party satisfies some exceptional circumstances. For example, the situations under which the party agreed the time-frame are such as were outside the reasonable contemplation of the parties, and that it would be just to extend the time, or the conduct of one party makes the agreed time-frame unjust,[50] or a substantial injustice would otherwise be done if the court does not extend the time.[51] Although the court possesses the power to extend the time-frame previously agreed by the parties in the aforementioned exceptional circumstances, recent cases concerning the Act show that

[41] Andrew Tweeddale and Keren Tweeddale, *Arbitration of Commercial Disputes: International and English Law and Practice* (OUP 2010), para 16.15.
[42] AA 1996, s 6.
[43] Blake, Browne, and Sime (n 36), para 31.05.
[44] ibid., para 31.07.
[45] Sutton, Gill, and Gearing (n 37), para 7-008.
[46] Blake, Browne, and Sime (n 36), paras 29.58-31.06.
[47] AA 1996, s 14.
[48] ibid., s 50.
[49] ibid., s 79.
[50] ibid., s 12(3).
[51] ibid., s 50.

the courts rarely extend time for commencing an arbitration when a party has missed a contractual time bar.[52]

3. The usual position is that arbitrators are appointed by the parties or through mechanisms agreed by them without any involvement of the courts.[53] In addition, there are default provisions in the non-mandatory category which permit the court to assist the parties in appointing the arbitrators. However, the mandatory provisions authorise the court to remove any preferred arbitrators of a particular party if at the instance of any aggrieved party the court is satisfied that there are justifiable doubts over the arbitrator's impartiality and qualifications.[54]
4. Mandatory provisions permit the court to intervene in determining 'any question as to the substantive jurisdiction of the tribunal.'[55] This provision is also a good example of the court's support to party autonomy because the court will not consider any such matter unless 'it is made with the agreement in writing of all the other parties…'[56]

The arbitration procedures depend fundamentally upon the agreement of the parties which causes a tension between the consensual basis of arbitration on the one hand and the establishment of an efficient arbitration 'system' on the other hand.[57] The Act (in particular mandatory provisions) may be viewed as a further step towards the creation of such a 'system.'[58] In other words, though the mandatory provisions have been enshrined in the Act as some restrictions upon party autonomy, those are, in true sense, conducive for a balanced nexus between party autonomy and judicial authority.

Rectifying Serious Injustice by Judicial Review

The court always retains an inherent power to intervene into any matter during exercising the power of judicial review notwithstanding any confinements specified in any law. Nevertheless, the Act defines some areas of confinements for the court. For example, the court may exercise its discretion while reviewing enforcement of an award,[59] challenges to the award on the ground of serious irregularity,[60] and appeals on a point of law.[61] During these reviews, does the Act confine the court within the mandatory provisions? Or does the Act give wide power to the court for reviewing any matters irrespective of mandatory or non-mandatory provisions for the sake of public interest?

At this stage, it is necessary to reiterate section 1(c) of the Act. It provides that, 'in matters governed by this Part the court should not intervene except as provided by this Part.' This provision itself allows the court to intervene into the matters of party autonomy by legislating

[52] *Cathiship SA v Allanasons Ltd* [1998] 2 Lloyd's Rep 511, [520] (Mr Geoffrey Brice QC).
[53] Blake, Browne, and Sime (n 36), para 31.15.
[54] AA 1996, s 24(1).
[55] ibid., s 32(1).
[56] ibid., s 32(2)(a).
[57] Karen Maxwell, 'English Arbitration Act 1996: Will Anything Change in Practice?' [1997] 13(4) Arbitration International 435.
[58] ibid.
[59] AA 1996, s 66.
[60] ibid., s 68.
[61] ibid., s 69.

'should' instead of 'shall.' In *AES- UstKamenogorsk v Ust-Kamenogorsk JSC,*[62] it was held that:

> *'[T]he use of the word 'should' in s.1(c) was also a deliberate departure from the more prescriptive 'shall' appearing in article 5 of the UNCITRAL Model Law. …in matters which might be regarded as falling within Pt 1 it is clear that s.1(c) implies a need for caution, rather than an absolute prohibition, before any court intervention.'*

The spirit of this principle invites the court's intervention into the matters decided following party autonomy. Consequently, it has been criticised for extending the court's power to treat an arbitral tribunal as an inferior branch of the judicial system.[63] It may be argued that party autonomy is restricted by section 1(c) regardless of the provision being mandatory or non-mandatory. However, questions may arise as to when the court's intervention might be allowed by the Act. It may be noted that such interventions would be allowed when the parties fail to follow the acceptable standard in their arbitration agreement.[64] As examples of this proposition of failure to follow the prescribed standards, the following discussions on sections 66, 68, and 69 of the Act are worthy of perusal.

Section 66 provides that an award made by the tribunal pursuant to an arbitration agreement may be enforced in the same manner as a judgement or order of the court to the same effect. Nevertheless, the court has discretion not to grant leave to enforce an award.[65] This discretion will be exercised in an appropriate case in the interests of justice and not as an administrative rubber stamping exercise.[66] In *Soleimany v Soleimany*,[67] a dispute between a father and a son was brought to the court where the plaintiff arranged the export of carpets from Iran in breach of the revenue laws and export controls of that country and the defendant sold the carpets in England and elsewhere. When disputes arose, both agreed for arbitration before the Beth Din in accordance with Jewish law. The Beth Din made an award in favour of the plaintiff, ignoring the issue of smuggling since it would have no effect under Jewish law. The plaintiff applied *ex parte* to have registered it as a judgement under English law. The English Court of Appeal stated:

> *'[W]here a foreign arbitration award was made pursuant to a valid arbitration agreement but was based on a contract which was illegal under the law of a friendly foreign state where that law governed the contract or the contract was to be performed in that state, the English court would not enforce that award on the grounds of public policy.'*

[62] [2013] UKSC 35, [33].
[63] Anthony Diamond, 'Publication Review on book *The Arbitration Act 1996: A Commentary*' [2004] 70(1) Arbitration 71.
[64] Fagbemi (n 6), 243.
[65] Sutton, Gill, and Gearing (n 37). para 8-005.
[66] *West Tankers Inc v Allianz Spa* [2012] EWCA Civ 27, [38].
[67] [1999] QB 785.

A liberal *dictum* than the *Soleimany* case came out in *Westacre Investments Inc v Jugoimport SDPR Holding Co Ltd*[68] where allegation of using personal influence and bribery were not considered as against public policy rather activities such as terrorism, drug trafficking, prostitution and paedophilia, corruption and fraud were considered as offensive to public policy. Hence, Wade[69] efficiently compared the findings of these two cases concluding that:

> *'[A] foreign arbitral award can be enforced even if the underlying contract offends against English public policy, providing that: (a) the award does not offend against any fundamental rule of English public policy; and (b) the award does not offend against the public policy of the governing law and/or the curial law; even if (c) the award is contrary to the public policy of the place of performance.'*

Considering the above precedent, it can briefly be said that the courts are not free to intervene into any arbitral award on the ground of violating public policy unless it contravenes any fundamental public policy which is 'necessary to safeguarding public interest.'[70] However, in reviewing any challenge for serious irregularity under section 68 of the Act, the court has the authority to scrutinise the arbitral tribunal's decision. In doing so, the court will enquire whether the tribunal has caused substantial injustice to the applicant and failed to follow the parties' agreement relating to its powers,[71] procedures,[72] issues in controversy[73] and form of the award.[74] It will also enquire whether the same has also been caused for non-compliance with core principles of justice such as violation of natural justice,[75] uncertain or ambiguous remedy,[76] attainment of award through fraudulent means or an award contrary to public policy,[77] and any other admitted irregularity.[78]

Regarding the tribunal's failure to adhere to the parties' agreement, the court should rectify the tribunal's mistake and also intervene where public interest requires it necessary for remedying substantial injustice in the proceedings. In *Lesotho Highlands Development Authority v. Impregilo SpA and Ors,*[79] it was alleged that the arbitrators exceeded their powers by expressing the award in European currencies and by awarding pre award interest in circumstances not permitted under Lesotho law. The House of Lords stressed on the necessity of focusing intensely on the particular power under the arbitration agreement but decided the matter as a mere error of law which would not amount to the Tribunal exceeding its power under section 68(2)(b). As mentioned earlier, the court and arbitral tribunal will follow the

[68] [2000] QB 288.
[69] Shai Wade, '*Westacre v Soleimany*: What Policy? Which Public?' [1999] 2 Int ALR 3 97.
[70] See discussion in part 2 of this article.
[71] AA 1996, s 68(2)(b and e).
[72] ibid., s 68(2)(c).
[73] ibid., s 68(2)(d).
[74] ibid., s 68(2)(h).
[75] ibid., s 68(2)(a).
[76] ibid., s 68(2)(f).
[77] ibid., s 68(2)(g).
[78] ibid., s 68(2)(i).
[79] *Lesotho Highlands Development Authority v Impregilo SpA and Ors* [2005] UKHL 43.

parties' agreement regarding determination of arbitral procedure. Nevertheless, there are examples of non-interference with the tribunal's decision which defied parties' subsequent agreement regarding the procedure because the court found it necessary for justice.[80] In fixing the issues of dispute, the court is not supposed to intervene into the parties' agreement nor tribunal's decision. Once it is recognised that a tribunal has 'dealt with' an issue, section 68(2)(d) does not allow any qualitative assessment as to how the tribunal dealt with it and it also does not matter whether it has done so well, inadequately, or indifferently.[81] However, the court reserves all authority to check whether failure to deal with any issue has caused substantial injustice.[82] For example, in *Secretary of State for the Home Department v Raytheon Systems Ltd,* the court directed to re-open an issue for conscious consideration.[83]

Similarly, section 68 is also embodied with some key judicial principles which are *sine quo non* for upholding justice for public interest. Section 68(2)(a) deals with matters 'such as bias, procedural unfairness and breach of natural justice.'[84] Violation of these principles must undergo judicial review regardless of the wide scope of party autonomy within the legal framework. In this regard, the judicial and arbitral principles of justice have been amalgamated as the test for bias is the same for both justices, jurors, and arbitrators.[85] However, the English law embraced the 'fair minded and informed observer' test whereby 'justifiable doubts' regarding the arbitrator's impartiality or independence[86] must be proved. Section 68(2)(g) allows the court to review the matters of party autonomy in case of any fraudulent practice causing substantial injustice. In case of an allegation of fraud practice, the Act allows the court to consider an innocent failure to give proper disclosure,[87] or the innocent production of false evidence[88] as non-fraudulent. Under the same provision, public policy is another ground of judicial intervention as mentioned in the discussion above with regards to section 66. An arbitral proceeding can also be challenged where any irregularity in the arbitral proceedings or award is decided as admitted by the tribunal or arbitral institution.[89]

An appeal on point of law under section 69 is restricted to party autonomy. The parties to arbitration may agree to exclude the appellate power of the court against any decision of the arbitral tribunal.[90] However, this restriction may be avoided through court's intervention for upholding public interest. This restriction means that appeals on questions of law are often 'dressed up' as challenges under section 68 which can be brought as of right.[91]

[80] *Secretary of State for Defence v Turner Estate Solutions Ltd* [2014] EWHC 244 (TCC).
[81] *Primera Maritime (Hellas) Ltd and others v Jiangsu Eastern Heavy Industry Co Ltd and another* [2013] EWHC 3066 (Comm), [40].
[82] *Petrochemical Industries Co (KSC) v Dow Chemical Co* [2013] 2 CLC 864 [15].
[83] Margarita N. Michael, 'Case Comment on Setting aside an award for serious irregularity: the Secretary of State for the Home Department v Raytheon Systems Limited' [2015] 18(2) Int ALR N13-N16.
[84] Harris, Panterose, and Tecks (n 18) para 68G.
[85] *AT&T Corporation v Saudi Cable Co* [2000] 2 Lloyd's Rep 127 [39].
[86] Austin I Pullé, 'Securing Natural Justice in Arbitration Proceedings' [2012] 20 (1) Asia Pacific Law Review 63, 85.
[87] *Profilati Italia SrL* v *PaineWebber Inc* [2001] CLC 672 [21].
[88] *Elektrim SA v Vivendi Universal SA* [2007] 1 CLC 16 [81].
[89] AA 1996, s 68(2)(i).
[90] ibid., s 69.
[91] Sutton, Gill, and Gearing (n 37), para 8-132.

The policy in favour of party autonomy does not permit derogation from the provisions of section 68. However, the matter of 'serious irregularity' specified in section 68 is tantamount as 'substantial injustice.' It must be followed preliminarily that any alleged irregularity has caused substantial injustice in case of any judicial intervention.[92] The test of 'substantial injustice' is intended to be applied by way of support for the arbitral process, not by way of interference with that process.[93] This discussion reveals that the court has the power to rectify any error caused by the arbitral tribunal, being subject to party autonomy. For example, the court during reviewing a tribunal's decision would consider whether the tribunal went beyond the parties' agreed power and proceedings. However, if the court finds the parties' agreement in contrary to public interest, party autonomy would not restrict the court to rectify it. Moreover, neither party autonomy nor the tribunal's discretion could refrain the court from exercising judicial review on the ground of fundamental principles of justice. This may be the reason for which the Departmental Advisory Committee on Arbitration Law stated that if the tribunal makes an error conducting an arbitration, it should be corrected for justice.[94]

Conclusion

The Arbitration Act 1996 adopted party autonomy and limited judicial intervention as its general principles and replaced some judicial authorities prevailed earlier through newly embraced mandatory provisions. This approach aims to avoid confrontation between the above two principles. This article preliminarily observes that party autonomy adopted in this Act is subject to two restrictions: the mandatory provisions of the Act, and the principles of public interest. Although this preliminary observation shows that the mandatory provisions are essentially restrictive for party autonomy, it appears that those were included in the Act with a view to using judicial powers to support an arbitration. Hence, the scheme of mandatory and non-mandatory provisions strikes a balance between the principles of party autonomy and limited judicial intervention. However, any autonomy provided to the parties is not beyond the court's scrutiny if it seems necessary for safeguarding public interests. The Act does not allow using the court as a rubber-stamping institution for legalising any party autonomy which is contrary to fundamental principles of justice and public policy. As an output of this cautious judicial approach, certain new principles, for example, 'substantial injustice', 'serious irregularity', 'fair minded and informed observer test' etc. have emerged in English jurisprudence. Thus, when the parties' agreement and tribunal's decision cause fundamental injustice, the Act would not restrict the court to intervene in an arbitration. It may be safe to conclude that allowing a broad extent of party autonomy while confining judicial intervention within the areas of fundamental injustice, the Act made the balance quite fair.

[92] *Lesotho* (n 79), para 28.
[93] Departmental Advisory Committee on Arbitration Law Report on the Arbitration Bill 1996, February 1996, para 280.
[94] *Petroships Pte Ltd v Petec Trading and Investment Corporation and Ors* [2001] 2 Lloyd's Rep 348.

Which Law is it Anyway? Comparative Approaches to Resolving the Applicable Law in Arbitration Agreements after *Kabab-Ji v Kout Food Group* [2021] UKSC 48

William Haslam

Abstract

Within two years, the UK Supreme Court gave judgement on two cases concerning the applicable law of arbitration agreements, arguably resolving over a decade of inconsistent decisions in the lower courts and abroad. Whilst the certainty brought by Enka Insaat Ve Sanayi AS v OOO Insurance Company Chubb [2020] UKSC 38 and Kabab-Ji v Kout Food Group [2021] UKSC 48 should be welcomed, it must not be detached from the international context in which commercial arbitration operates. This case note will evaluate the recent UK Supreme Court's decision in Kabab-Ji and contrast it with the diverging approach in the same case in the French Court of Appeal. A clear tension can be identified between these cases concerning how willing courts are to depart from the express intention of the parties over the necessary implication of the chosen law. Moreover, this comparative review evidences the mosaic influence of domestic apex courts in international arbitration. Such courts contribute to a palimpsest of conflicting authorities concerning a fundamental question of any arbitration to be determined by the tribunal: what is the applicable law of the arbitration agreement? By taking an international perspective, this case note identifies how a path of clarity can be traced in a previously unsettled area of arbitration law.

Introduction

The question of deciding the governing law of an arbitration agreement[1] in the absence of express indication, is a question that continues to divide the international arbitration community. However, from a UK perspective, Popplewell LJ's call to 'impose some order and clarity' on this area of law now seems to have been realised.[2] This choice of law question falls outside the scope of the Rome I Regulation,[3] and is instead determined by the common law rules based on Article V(I)(a) of the New York Convention.[4] In answering this question, two competing approaches can be identified that differ in the relative weight given to the curial law[5] versus the substantive law of the agreement. The first, favoured by the UK courts,[6] places

[1] The law governing the arbitration agreement is distinct from the '*lex arbitri'* or 'curial law.' These laws govern the process of the arbitration itself, not the validity and scope of the arbitration agreement. Redfern & Hunter, *International Arbitration* (6th edn, OUP 2016), 161.
[2] *Enka Insaat Ve Sanayi AS v OOO Insurance CO Chubb* [2020] EWCA Civ 574, [89].
[3] Regulation (EC) 593/2008 art 1(2)(e).
[4] Convention on the Recognition and Enforcement of Foreign Arbitral Awards 1958 ('the New York Convention').
[5] The curial law (or '*lex arbitri'*) is the law which governs the process of an arbitration. This is determined by the parties' choice of an arbitral seat. See Dicey, Morris, and Collins, *The Conflict of Laws* (15th edn, Sweet & Maxwell), paras 16-35.
[6] *Sulamerica Cia Nacional de Seguros SA v Enesa Engenharia SA* [2012] EWCA Civ 638.

greater emphasis on an express choice of substantive law to be interpreted as an implied choice of law governing the arbitration agreement. Against this, a growing international body of decisions,[7] driven particularly by the French courts,[8] emphasises the unique character of arbitration clauses which should not so readily follow the substantive choice of law of the contract. Instead, this line of cases suggests the choice of the seat should be determinative.

In part, this fault line can be explained through differing perspectives to the inherent nature of arbitration clauses and the varying weights given to the doctrine of separability. The former approach sees arbitration clauses as merely another term of the contract, with no *sui generis* character which allows departure from an express choice of law. The latter rejects this approach. It fails to acknowledge the unique role of dispute resolution clauses and instead imposes on the parties an arbitral law which was not expressly agreed by them. Why should an *express* choice of arbitral seat be displaced by an implied choice of law? Particularly as there are strong practical and policy reasons why the *lex arbitri* should match the law of the seat. For example, to allow courts of the seat jurisdiction to apply domestic law to assist the arbitral process. Consequently, as Lord Hamblen and Lord Leggatt aptly stated in *Kabab-Ji,* 'there is nothing approaching a consensus'[9] on the question of which law should fill in if no law is specified in an arbitration agreement. Yet, the twin judgments of *Enka Insaat Ve Sanayi AS v OOO Insurance Company Chubb*[10] and *Kabab-Ji*[11] provide considerable clarity on where, at least, consensus is to be found in the UK Courts.

Kabab-Ji in the UK Supreme Court

Whereas *Enka* dealt with the question of arbitral law before the commencement of arbitration, *Kabab-Ji* addressed the same question at the enforcement stage. Unsurprisingly, the court quickly emphasised that it would be 'illogical'[12] to address the question of arbitration differently depending on the stage of arbitration, signalling the unity between these two judgments.

In brief, the facts of *Kabab-Ji* concerned a Franchise Development Agreement (FDA) signed in 2001 between the claimant and Al Homaizi Group (AHG) for the franchising of a Lebanese restaurant chain across Kuwait. Under Clause 15, the FDA was expressly governed by English Law. In 2005, AHG underwent a corporate restructuring to become a subsidiary of Kout Food Group (KFG), although the FDA was not updated to reflect this fact. When a dispute later arose under the FDA, it was submitted to ICC Paris seated arbitration as expressly required by Clause 14 of the FDA.

The central question was whether KFG had become a party to the FDA and was therefore bound by the arbitration clause. KFG argued it was not. KFG nevertheless participated in the arbitration under protest and, when the tribunal issued an award of $6.6 million against them, appealed to the English Courts to have the award annulled on the basis they were not a party to the arbitration agreement.

[7] *AT&T Mobility LLC v Concepcion* 131 S.Ct 1740, 1753 (2011).
[8] *Municipalité de Khoms El Mergeb v Société Dalico,* 1994 Rev arb 116, 117 (French Cour de cassation civ 1e).
[9] *Kabab-Ji v Kout Food Group* [2021] UKSC 48, [32] (Lord Hamblen and Lord Leggatt).
[10] [2020] UKSC 38.
[11] *Kabab-Ji* (n 9).
[12] ibid., [35].

The dispute centred around what law should govern the procedure of the arbitration agreement. Kabab-Ji contended, as agreed by the ICC arbitral tribunal and later the Paris Court of Appeal, that the express selection of Paris seated arbitration indicated French law was the curial law.

Yet, the UK Supreme Court, confirming the judgement of the Court of Appeal, reached the opposite conclusion. The law governing the arbitration agreement was English law which was found to be the same as the substantive contract. Further, under English law, KFG had not been added to the FDA (there was no 'novation by addition' as *Kabab-Ji* attempted to argue)[13] and therefore could not be bound by the arbitral award.

How can these approaches be reconciled? In short, not easily. This case note suggests that the cause of this division is two-fold. First, different views towards the sanctity of the arbitration agreement and second, the differing weights given to the importance of business efficacy. It is also worth noting that an outstanding appeal to the French Cour de Cassation may yet unify these conflicting views, although, as will become clear, this seems unlikely.

The UK court's judgement in *Kabab-Ji* signals a doubling down of the *Enka* approach which favours a pragmatic, if formalistic, resolution of curial law uncertainty. It is common ground that the starting point for determining the law of the arbitration agreement is the ambiguously worded Article V(1)(a) of the Convention on the Recognition and Enforcement of Foreign Arbitral Awards 1958 (the New York Convention). This provision provides that arbitration agreements are governed by either (i) the law the parties have subjected it to or, 'failing any indication thereon', (ii) the law of the country in which the award was made (which, generally, is the law of the seat). In both *Enka* and *Kabab-Ji,* the UK Supreme Court is reluctant to fall back on the second limb of Article V(1)(a) by interpreting that the condition precedent to access this 'default' choice of law is a high bar. Specifically, there must not be 'any indication' of the parties' preference for the arbitral law and such indication does not need to be express or even specific.[14]

When expressed in this way, it becomes difficult to displace the presumption that the general choice of substantive law will 'normally be sufficient indication,'[15] preventing any recourse to the default choice of law. This reluctance may be with good reason. It prevents the counter-intuitive situation that, despite an 'entire agreement' clause (as was the case in *Kabab-Ji*),[16] arbitration clauses are somehow impliedly excluded from declarations as to the governing law of the contract, as expressly agreed by the parties purely because of their *sui generis* character as dispute resolution clauses. For the English courts, any deviation from this express choice of law is a step too far. In other words, *any* express choice - regardless of whether directed at the arbitration agreement or not - is a sufficient 'indication' to apply both to the substantive contract and the arbitration agreement. Therefore, once the first limb of Article V(1)(a) is engaged (as will often be the case with any choice of law clause), any question of the curial law following that of the seat quickly falls away.

[13] ibid., [62].
[14] ibid., [33].
[15] ibid., [35].
[16] ibid., [39].

The reasoning in *Kabab-Ji* rests on the premise that English law is reluctant to acknowledge any special character of arbitration clauses. This is consistent with Born's analysis of the English approach which is reluctance to embrace 'sweeping formulations of any general principles of "autonomy" or "independence"'[17]. *Kabab-Ji* can be read as a further example of such reluctance. Instead, the express intention of the parties is foregrounded to the extent that no room is left for an implied choice of law to govern the arbitration agreement.

Yet, by placing this premise in the context of international commercial arbitration, it can be suggested that the UK Supreme Court too readily extends an express choice of law to a unique and important clause. By denying this *sui generis* status, the UK Supreme Court is subtly side-stepping the question of separability and instead using the broad provision of 'express intent of the parties' to gloss over the difficult questions about a free-standing arbitration agreement. What makes this position more intriguing is that English courts readily embrace separability when this is required to ensure the survival of the arbitration clause from an invalid underlying contract,[18] yet fails to place the same weight on the sanctity of the dispute resolution clause for the purposes of determining an ambiguous curial law. This reluctance to fully embrace separability is in contrast with the international acceptance of the doctrine as a 'conceptual and practical cornerstone'[19] of international commercial arbitration and the approach of other apex courts.

Kabab-Ji in the Paris Court of Appeal

In contrast, the Paris Court of Appeal adopted an approach of whole-hearted separability that was not impacted by the 'whole agreement' clause in the FDA. Before even turning to contractual interpretation, the French court plainly stated that 'pursuant to a substantive rule of international arbitration law, the arbitration clause is legally independent from the underlying contract.'[20] This would not be accepted by the UK Supreme Court. Alongside a 'legally independent' character comes an expectation that parties always have two choices of law to make (substantive law and curial law) and, crucially, there is no presumed connection between these two choices. For the Paris Court, these matters are siloed, while for the UK Supreme Court these are closely connected.

Moreover, the Paris Court places considerable weight on the absence of any express provision for English law to govern the arbitration agreement,[21] and the absence of any implied evidence that establishes 'unequivocally the common will of the parties to designate English law' as governing the validity, transfer, and extension of the arbitration clause. It is interesting to note here that the Paris Court does not even consider the evidential weight of the clause regarding choice of law, reaffirming that the arbitration clause is a distinct matter which cannot be readily implied. It is submitted that this approach adopts a more realistic interpretation of Article V(I)(a) of the New York Convention which needs the 'indication' required to specifically relate to the arbitration clause, thereby placing more weight on the choice of seat.

[17] Gary Born, *International Commercial Arbitration* (3rd edn, Wolters Kluwer, 2020), 405.
[18] *Fiona Trust & Holding Corp. v Privalov* [2007] EWCA Civ 20 [29], affirmed [2007] UKHL 40; *Lesotho Highlands Development Authority v Impregilo SpA* [2006] 1 AC 221, 232.
[19] Born (n 17), 376.
[20] *Kabab-Ji SAL (Lebanon) v Kout Food Group (Kuwait)* [CA Paris, 23 June 2020, no. 17/22943], [25].
[21] ibid., [29].

Such emphasis is logical between the law governing the underlying contract and the choice of seat, it is surely the latter which most closely elucidates the parties' intention regarding how arbitration should be conducted. It is important to note that the Paris Court of Appeal is not relying on the fall-back 'default' choice of law under Article V(I)(a) but is able to reach the same conclusion with reference to the intention of the parties. However, unlike the English courts, that question of implied intention is asked narrowly and based on the starting point of an arbitration agreement's complete free-standing legal status. The English courts do not presume such a starting point although, as discussed above, will do so if the validity of the underlying contract is imperilled. Consequently, there is a satisfying logical coherence to the Paris Court's approach if (and this is a decisive 'if' for the English courts) the *sui generis* status of arbitration agreements is accepted.

Finally, there is practical coherence to the Paris Court's seat-orientated approach. Since curial law concerns the validity of the arbitration agreement, it makes practical sense that the court in the jurisdiction of the seat is the one to determine such an important question of validity of the arbitration agreement. Evidently, the international character of arbitration does not prevent a transnational relationship between the court enforcing curial law and the tribunal, but it seems contrary to the harmonised policy of selecting an arbitral seat in the first place in which the judicial reputation of that state so heavily influences.

Conclusion

Overall, the English Supreme Court's decision in *Kabab-Ji* solidifies the position in *Enka* whilst exposing English law's reluctance to fully embrace the free-standing status of arbitration clauses. It may be argued that Born's analysis suggesting English law's approach of being reluctant to embrace 'sweeping formulations of any general principles of 'autonomy' or 'independence' may be a stronger contention than the other contention discussed in this article.

This is only accentuated by the contrasting Paris Court of Appeal judgement which, at the very least, suggests the ongoing international curial law debate is not as settled as the Supreme Court may have hoped (particularly following the quick succession of *Enka* and *Kabab-Ji*). *Kabab-Ji* attempts to resolve this debate but instead draws attention to a division in the doctrine of separability. Whilst English courts have long paid lip-service to a fully-fledged separability, *Kabab-Ji* invites us to question the strength of this belief.

Finally, an interesting sub-text to the UK Supreme Court judgement is a clear dissatisfaction with the lack of explanation of the arbitral tribunal's reasoning, particularly concerning whether the law of the contract or seat should apply to the arbitration agreement.[22] In many ways the same criticism can be levelled at the Paris Court of Appeal decision that relies heavily on 'generally recognised principles of law',[23] and deals with the question merely within sixteen paragraphs. This perhaps evidences the tension between both civil and common law jurisdictions and, more fundamentally, between the law as developed through private-arbitration proceedings and the detailed analysis of apex courts. It is this tapestry of decision

[22] *Kabab-Ji* (n 9), [46].
[23] *Kabab-Ji* (n 20), [25] and [29].

making which allows the debate around determining ambiguous curial law to continue, despite the best efforts of the UK Supreme Court.

The Efficacy of the International Criminal Court: The Relationship Between the Crime of Aggression and Deterrence

Cham Mikhael

Introduction

Domestic criminal law plays a vital role in deterring perpetrators from committing heinous acts. The International Criminal Court (ICC), being the only permanent criminal justice system in the international sphere, was established to mimic that deterring effect for the most serious crimes. It is argued by many scholars that if such systems 'set the costs of crime sufficiently high, potential criminals will rationally choose not to violate the law.'[1] Criminological theory presents us with two elements that set such costs – the certainty of punishment and the severity of the sentence.[2] Nevertheless, these costs are only proven useful when criminals believe that the regime applies to them and they change their behaviour in response.[3] Thus, an objective survey on the cost-benefit calculation which gives rise to deterrence is tempered with the subjectivity of the criminal's interpretation.

A potential hurdle in ascertaining deterrence in the international sphere is the difficulty of determining the link between the persons of authority, or moral agents, and the community it serves to protect – this, instead, is clear in domestic criminal law systems. Duff recognises the importance of citizenship in domestic law as a key component to ascertaining that relationship.[4] Transplanting this unto international criminal law, the relevant community is simply the moral community of humanity, a rather intangible link compared to citizenship.[5] Another hurdle for the importation of deterrence from domestic criminal justice systems into international criminal law is the notion that criminals that commit the most serious crimes under international criminal law are 'profoundly irrational and motivated by bloodlust, religious fervour or ancient ethnic hatred.'[6] Proceeding on this view halts any discussion regarding the deterrence of aggression since it places these criminals on a higher level of criminal hierarchy whereby, logically, standard deterrence strategies are futile. Having said that, by assessing the benefits and gratifications derived from the commission of mass atrocities, judges may impose serious punishment that counter those benefits and significantly add to the deterring influences in international criminal law.

[1] Kate Cronin-Furman, 'Managing Expectations: International Criminal Trials and the Prospects for Deterrence of Mass Atrocity' [2013] 7(3) The International Journal of Transitional Justice 434, 439.
[2] ibid., 441.
[3] Andreas von Hirsch and others, *Criminal Deterrence and Sentencing Severity* (Hart Publishing 1999).
[4] Anthony Duff, 'Responsibility, Citizenship, and Criminal Law', in RA Duff and Stuart Green (eds), *Philosophical Foundations of Criminal Law* (OUP 2011), 125-48.
[5] Ryan Long, 'Responsibility, Authority, and the Community of Moral Agents in Domestic and International Criminal Law' [2014] 14(4-5) International Criminal Law Review 836, 837.
[6] Cronin-Furman (n 1).

In reality, the significance of deterrence in the international community emerged quite recently, as previous tribunals were more focused on delivering justice through retribution.[7] The scope of justice expanded to include deterrence when ad hoc tribunals (which were established *during* an ongoing conflict) were able to curtail the violence.[8] Though these tribunals were created with a punitive aim and were limited to the conflicts at hand, a joint statement by these ad hoc tribunals asserted that 'only a sustained commitment to accountability will deter atrocities.'[9] In fact, the judgement in the *Kambanda* case alluded that deterring crime, ranks above retribution.[10]

The Rome Statute of the International Criminal Court cemented such intolerance to mass atrocities by aiming 'to put an end to impunity for the perpetrators of these crimes and thus to contribute to the prevention of such crimes.'[11] Nonetheless, the notion that increasing the number of trials for mass atrocities or creating a permanent system of accountability would dissuade criminals is not supported by empirical evidence.[12] The link between international legal accountability and general deterrence is muddied by geographical and temporal limitations and the existence of other post-conflict institutions.[13] Even though the ICC tries to change this, the aforementioned empirical realities coupled with the procedural and fundamental difficulties found in prosecuting acts of aggression, renders ineffective deterrence of the crime.[14] Furthermore, the incentive for state leaders to commit aggression, especially when sugar-coated with a misleading vision of justice that is widely supported, outweighs the slim chances of prosecution and punishment set out by the ICC's narrow set of conditions and jurisdiction. This work explores the ambiguity of the definition of aggression, the obstacles found in its implementation and lastly, the lack of reinforcement, finally coming to the conclusion that the ICC's ability to set sufficiently high costs for the crime is unclear.

Background: Defining Aggression and the Complications of the Rome Statute

Commitment to outlawing war under international law can be traced back to the Covenant of the League of Nations in 1919 and then subsequently to the Kellogg-Briand Pact in 1929. After the Second World War, the Charter of the United Nations embedded such condemnation in Article 2.4 by 'prohibiting the threat or use of force and calls on all Members to respect the

[7] *France and ors v Göring (Hermann) and ors* (1946) 41 AJIL 172 <https://www.loc.gov/rr/frd/Military_Law/pdf/NT_Vol-I.pdf> accessed 23 January 2022.
[8] See International Criminal Tribunals for the former Yugoslavia and Rwanda (ICTY and ICTR respectively).
[9] University Of Minnesota, Human Righst Library, *Joint Statement of the Prosecutors of the International Criminal Court, the International Criminal Tribunal for the former Yugoslavia, the International Criminal Tribunal for Rwanda, and the Special Court for Sierra Leone* (27 November 2004) <http://hrlibrary.umn.edu/intribjointstatement2004.html> accessed 30 December 2020.
[10] *Prosecutor v Kambanda*, Judgement and Sentence ICTR-97-23-S, The International Criminal Tribunal for Rwanda, 4 September 1998, [28].
[11] Preamble, Rome Statute of the International Criminal Court, opened for signature 17 July 1998, 2187 UNTS 3 (entered into force 1 July 2002) (Hereinafter Rome Statute).
[12] Cronin-Furman (n 1) 435.
[13] ibid., 440.
[14] Oskar NT Thoms, James Ron and Roland Paris, 'The Effects of Transitional Justice Mechanisms: A Summary of Empirical Research Findings and Implications for Analysts and Practitioners' [2008] Ottawa: Centre for International Policy Studies, University of Ottawa <http://aix1.uottawa.ca/~rparis/CIPS_Transitional_Justice_April2008.pdf> accessed 30 December 2020.

sovereignty, territorial integrity and political independence of other States'[15] – except for individual or collective defence and authorised use of force by the UN Security Council.[16] The definition of aggression today manifests the principles of the aforementioned section of the UN Charter, alongside the tribunal judgement and resolution mentioned below.

Aggression made its debut as a punishable international crime in the International Military Tribunal (IMT) at Nuremberg as a 'crime against peace', and was dubbed the 'supreme international crime' because it 'provides the occasion for the commission of other crimes.'[17] This understanding was then adopted in UN General Assembly Resolution 3314, serving as a non-binding guide to the UN Security Council on the definition of aggression but not on individual conduct.[18] In 1998, the Rome Statute tried embracing this legacy by mentioning the crime of aggression in principle, but countries could not decide on a definition or jurisdiction, thus deferring it for future revision. From the outset, there was a fundamental dispute over the crime of aggression. Since then, this crime has been extremely politicised and there has been disagreement over what constitutes a criminal act of aggression since it could not be resolved.[19] The 2010 Review Conference in Kampala passed a resolution that amended the Rome Statute to include a definition for aggression, a jurisdiction and a set of trigger mechanisms; but, its commencement was delayed again due controversy surrounding the idea of criminalising aggressive war.[20] After much compromise, 30 states ratified the Kampala amendments, and the Assembly of State Parties later activated ICC jurisdiction in 2018.[21]

Article 8*bis* states that aggression is the commission by state leaders to direct military or political action of a state for unlawful purposes onto another state.[22] A potential drawback of focusing on a political or military leader is the 'relief' of other senior officials not under this definition. Since waging war in a technical sense is now uncommon, aggression was drafted to amount to a 'manifest violation' of the UN Charter 'by [the act's] character, gravity and scale.' Though it is certain a vague future act does not constitute aggression, scholars and lawyers remain unsure on whether it involves all the three elements or if some take more priority than

[15] Charter of the United Nations [1945] 1 UNTS XVI (UN Charter) art 2(4) <https://treaties.un.org/doc/publication/ctc/uncharter.pdf> accessed 23 January 2022.
[16] ibid., arts 51 and 42.
[17] Robert Cryer, Darryl Robinson and Sergey Vasiliev, *An Introduction to International Criminal Law and Procedure* (4th edn, Cambridge University Press 2019), 298.
[18] UNGA Res 3314 (XXIX) (14 December 1974).
[19] Zachary Manfredi and Julie Veroff, Exclusive Interview with Eric Leonard, Professor of Political Science, Shenandoah University, Oxford Transitional Justice Research Debating International Justice in Africa (OTJR Collected Essays, 2008-2010), 170 <https://www.law.ox.ac.uk/sites/files/oxlaw/justice_in_africa1.pdf> accessed 30 December 2020.
[20] Douglas Guillofyle, 'The Crime of Aggression' (24 August 2018) <https://www.youtube.com/watch?v=0BKVU1vM50U> accessed 30 December 2020.
[21] Rome Statute of the International Criminal Court, arts 8*bis*, 15*bis*, 15*ter* and 25(3*bis*).
[22] Full definition of the crime of aggression is set out in paragraph 1 and 2 of Article 8*bis* '1. For this Statute, 'crime of aggression' means the planning, preparation, initiation or execution, by a person in a position effectively to exercise control over or to direct the political or military action of a State, of an act of aggression which, by its character, gravity and scale, constitutes a manifest violation of the Charter of the United Nations. 2. For paragraph 1, 'act of aggression' means the use of armed force by a State against the sovereignty, territorial integrity or political independence of another State, or in any other manner inconsistent with the Charter of the United Nations. Any of the following acts, regardless of a declaration of war, shall, in accordance with United General Assembly resolution 3314 (XXIX) of 14 December 1974, qualify as an act of aggression […].'

the rest. This threshold is also inextricably linked to the collective act of one state to another, thus ruling out non-state aggressors or victims in this definition.

The definition of aggression is still a contentious matter with varying interpretations on what unlawful aggressive acts may be. Since it has never been tried in the ICC, it is unclear whether acts with a moral reason, such as humanitarian intervention or dealing with a threat in foreign territory, are part of the 8*bis* definition.[23] Be that as it may, scholars have also expressed a pertinent concern, in that the risk of prosecution by the ICC could actually deter political and military leaders from launching military intervention that serve a legitimate goal and promote community interests.[24] The crime of aggression's jurisdictional regime discussed below further discourages countries from ratifying the Kampala amendments.

Jurisdictional Limitations and its Effects

As of 2020, only 40 countries have ratified the Kampala amendments.[25] As long as militaries around the world continue to use force in situations that leaders claim are defensive or humanitarian, a legal regime that tests such uses of force is unwelcome.[26] Furthermore, the criticisms faced by the making and functioning of the ICC itself contribute to the argument that there is no advantage to adopting the Kampala amendments.

The exercise of ICC jurisdiction over the crime of aggression is found in articles 15*bis* and 15*ter*. Generally, both the victim and aggressor states must ratify the amendment and *not* opt out from the ICC's exercise of jurisdiction. This opt-out provision complicates the ICC's jurisdictional regime. Moreover, the only instances where jurisdiction is certainly applicable is when the *aggressor* state has both ratified and not opted out – placing the significance on the aggressor rather than the victim. There are also diverging interpretations on the exercise of jurisdiction when faced with an aggressor state that has not ratified *but* not opted out.[27] The circle of countries included in the ICC's jurisdiction further shrinks when it is acknowledged that non-ICC states, regardless of being victims or aggressors, are exempt from prosecution (except in Security Council referrals).

In instances other than UN Security Council referral, the UN Security Council must still make a determination of the proposed act of aggression (either by state referral or *proprio motu*). If a determination is not given within 6 months, the Pre-Trial Division of the ICC may authorise an investigation. Though this operation increases the Prosecutor's preparedness in bringing the charge to court, this lengthy process in-fact delays the operations of the ICC and the delivery of justice. Moreover, the court can only exercise its jurisdiction one year after a

[23] Harold Hongju Koh and Todd F. Buchwald, 'The Crime of Aggression: The United States Perspective' [2015] 109 The American Journal of International Law 2, 257.
[24] Dapo Akande, 'What Exactly was Agreed in Kampala on the Crime of Aggression?' (*Blog of the European Journal of International Law*, 21 June 2010) <https://www.ejiltalk.org/what-exactly-was-agreed-in-kampala-on-the-crime-of-aggression/> accessed 1 January 2021.
[25] Amendments on the crime of aggression to the Rome Statute of the International Criminal Court, united Nations, Kampala as available on <https://treaties.un.org/Pages/ViewDetails.aspx?src=TREATY&mtdsg_no=XVIII-10-b&chapter=18&lang=en> accessed 30 December 2020.
[26] Benjamin B Ferencz, 'Ending Impunity for the Crime of Aggression' (2009) 41 Case Western Reserve Journal of International Law 2, 281.
[27] Coalition for the International Criminal Court, 'The Crime of Aggression' https://www.coalitionfortheicc.org/explore/icc-crimes/crime-aggression> accessed 1 January 2021.

state ratifies the Kampala amendments, further prolonging justice and accountability. However, advocating for an expansion of ICC jurisdiction comes with its own hurdles. Countries are guarded with many protective layers intending to preserve the sovereignty of the nation and the rights of its people. Investigating perpetrators in a country that has not given its consent would infiltrate such protective layers and discredit the ICC.

Furthermore, the principle of complementarity – which 'represents the idea that states, rather than the ICC, will have priority in proceeding with cases' – is discordant with the nature of aggression.[28] Except when the aggressor states are defeated, it is unlikely that they would prosecute their own leader, let alone cooperate with the ICC in gathering evidence. This cooperation can also be constrained by power politics when the ICC investigates a case and an arrest warrant need implementing.[29] The activation of aggression risks places the ICC in politically sensitive situations, making controversial decisions that may undermine the legitimacy of the ICC which must be free from bias. Overuse of the crime of aggression can also hinder the Court's power, especially when perpetrators are acquitted.[30] If, as an alternative, the ICC prosecutes perpetrators for commission of genocide, war crimes or crimes against humanity in place of the crime of aggression, it would have to explain why it is not acting in relation to aggression and squander the decades-long efforts by the international community to finally bring the crime of aggression to fruition.[31] As a result, failing to successfully prosecute perpetrators for aggression would embolden them to commit further atrocities and eliminate any form of deterrence set out by the international community.

UN Security Council Implication

Article 13(b) of the Rome Statute allows the UN Security Council to refer situations to the Prosecutor acting under Chapter VII of the UN Charter. Notwithstanding the gridlocked opposition of its permanent members, the augmented role of the UN Security Council in determining whether crimes have taken place 'risks supplanting the independent investigations and role of the Prosecutor as well as the apolitical vision of the judicial process.'[32]

It is suggested that to strengthen the condemnation of the crime of aggression, regular reports and public statements on the proliferation of aggression must be issued by the Prosecutor to pressure the Security Council and ICC parties to refer more cases. It would display both the ICC's commitment holding perpetrators accountable and the intention to deter

[28] Linda E Carter, 'The Principle of Complementarity and the International Criminal Court: The Role of Ne Bis in Idem' [2010] 8 Santa Clara Journal of International Law 165.
[29] Catherine Gegout, 'The International Criminal Court: limits, potential and conditions for the promotion of justice and peace' [2013] Third World Quarterly <https://www.tandfonline.com/doi/full/10.1080/01436597.2013.800737> accessed 1 January 2021.
[30] Manfredi and Veroff, Exclusive Interview with Eric Leonard (n 19), 171.
[31] Carsten Stahn, *A Critical Introduction to International Criminal Law,* (Cambridge University Press 2018), 105.
[32] Coalition for the International Criminal Court, 'Factsheet: The Crime of Aggression Within the Rome Statute of the International Criminal Court' <https://www.coalitionfortheicc.org/sites/default/files/cicc_documents/CICC-%20Factsheet%20Crime%20of%20Aggression%20Final-%20changes%2027Nov2019.pdf> accessed 2 January 2021.

the escalation of further aggression. This nevertheless casts doubt on whether these further actions by the Prosecutor would be encroaching on the role of the UN Security Council.[33]

Absence of Law Enforcement

Ever since the establishment of international criminal tribunals, prosecutors face the insurmountable challenge of apprehending indictees. Banteka furthers this by emphasising that the success of the ICC solely depends on state cooperation and compliance regarding the arrest and surrender of suspects.[34] The principal difference between domestic criminal systems and international criminal regimes is the lack of police to enforce the law. The closest international enforcement body is the UN Peacekeeping, but even it has its limits. There is no police force to arrest perpetrators of mass atrocities and the ICC must rely on states and their cooperation to do the work on the ground.[35]

The presence of the police provides practical assistance to deterring citizens from committing crimes condemned by the domestic courts. Although creating an extensive, international enforcement agency would assuredly discourage criminals from committing crimes of aggression, it would be practically impossible to establish such an agency given that the ICC already operates on such a narrow scale. Unfortunately, no other method of enforcement can fill that void in international criminal accountability to mimic domestic police forces, leaving perpetrators feeling less threatened by the international criminal regime.

Banteka associates the hurdle of successful arrests to the ICC's 'delusive persistence in rejecting the factoring of politics in their operation.'[36] As previously mentioned, maintaining impartiality is one of the main objectives while investigating and prosecuting; there is a risk that politically sensitive situations give rise to controversial decisions which undermine the legitimacy of the ICC. Banteka counters this by instead placing politics at the fore of the ICC's role – calling for a capitalisation of its inherent political role.[37] This could be an alternative to the current enforcement regime, allowing the Prosecutor to improve compliance with ICC arrest warrants through external pressure through positive (membership and development aid)[38] and negative (travel bans and asset freezes)[39] inducements. Such inducement may arise from ICC members, as well as other actors, such as NGOs and transnational networks. Indeed, this may be a regime that enhances the efficacy of enforcement in the international sphere.

Conclusion

Without the power to enforce justice, the crime of aggression in the Rome Statute stands as a symbolic allusion to the Nuremberg Principles. Once the ICC can bypass technical and bureaucratic impediments, it can prosecute perpetrators for the crime of aggression but yet be faced with extremely controversial questions on the use of force that have puzzled the

[33] Office of the Prosecutor, 'Invited Experts on Aggression Question' (ICC Forum 2018) <https://iccforum.com/aggression> accessed 2 January 2021.
[34] Nadia Banteka, 'Mind the Gap: A Systematic Approach to the International Criminal Court's Arrest Warrants Enforcement Problem' [2016] 49 Cornell International Law Journal 528.
[35] Article 86 of the Rome Statute sets out the general obligation to cooperate with the Court.
[36] Banteka (n 34), 521-524.
[37] ibid., 533.
[38] ibid., 539.
[39] ibid., 541.

international community for decades. Until then, the absence of convictions on aggression, coupled with limits on individual criminal responsibility, make it difficult to ascertain the ICC's success in deterring individuals from committing the crime of aggression.

Bridging the Void: The Potential of Recent Supreme Court Decisions to Curtail Supply Chain Human Rights Abuses

Jack Palmer*

Abstract

Changes in the global economy have produced regulatory black holes where human rights abuses are prevalent beyond the purview of effective legal jurisdiction.[1] In the past three decades, countries at the consumption end of global supply chains have become increasingly cognisant of this systematic lacuna and the abuses of which it is productive.[2] Accordingly, numerous potential legislative, diplomatic and corporate self-governance developments have been mooted.[3] However, it is via the common law that the UK Supreme Court, in the cases of Vedanta and Okpabi, has expanded the potential for victims of foreign human rights abuses to litigate UK domiciled parent companies for the conduct of their subsidiaries. This article explores the extent to which these decisions might form the basis for a new mode of human rights litigation, the limitations thereof and the likely implications for courts, businesses, and victims of human rights abuses.

Introduction

'One of the problems with globalisation is that the law lags behind commerce.' [4]

Changes in the global economy have necessitated accordant changes in the regulatory frameworks designed to reduce human rights abuses. The global economy has increasingly become dominated by modes of production that seek fluid transfers of capital and the export of production to areas of maximal productivity.[5] In this context, the English and Welsh Legal system has been ill-equipped to curb the human rights abuses that, in the absence of an effective legal framework for liability, this organisation of economic production perpetuates.[6] This

[1] Human Rights Watch, 'Human Rights in Supply Chains' [2016] <https://www.hrw.org/report/2016/05/30/human-rights-supply-chains/call-binding-global-standard-due-diligence> accessed 8 January 2022.
[2] Rhys Jenkins, 'Codes of Conduct: Self-Regulation in the Global Economy' UN Research Institute for Social Development: Technology, Business and Society Programme, Paper No. 2, April 2001) < https://www.researchgate.net/publication/37150822_Codes_of_Conduct_Self_Regulation_in_a_Global_Economy> accessed 8 January 2022.
[3] Carolijn Terwindt and others, 'Supply chain liability: pushing the boundaries of the common law?' [2018] 8 Journal of European Tort Law 261.
[4] Ian Binnie, 'Confronting Corporate Complicity in International Human Rights Abuses' (Canadian Bar Association, 6 May 2010) <http://www.cba.org/cba/cle/PDF/INTL10_Binnie_Speakingnotes.pdf> accessed 30 October 2021.
[55] Donatella Alessandrini, 'The Time that Binds the 'Trade-Development' Nexus in International Economic Law' [2001] 12 Trade, Law and Development 625.
[6] Sheldon L Leader and Anil Yilmaz-Vastardis, 'Improving Paths to Business Accountability for Human Rights Abuses: A Legal Guide' [2017] Essex Business and Human Rights Project <https://www1.essex.ac.uk/ebhr/documents/Improving-Paths-to-Accountability-for-Human%20Rights-Abuses-in-the-Global-Supply-chains-A-Legal-Guide.pdf> accessed 22 October 2021; Thomas Clarke and Martin

article will illustrate the disconnect between regulatory oversight and this globalised mode of production and how this disconnect is productive of human rights abuses. Subsequently, it will weigh the recent Supreme Court decisions of *Vedanta*[7] *and Okpabi*[8] and the potential of these cases to partially redress the regulatory void. In doing so the article will make two points; firstly, *Vedanta* and *Okpabi* are representative of a transition in favour of foreign direct liability which will open new avenues for victims of human rights abuses to directly pursue redress in the English and Welsh legal system. Secondly, this potential is fundamentally limited in a number of ways meaning these cases are not a panacea for human rights abuses. The article will conclude by contextualising the likely effects of these cases for victims, businesses and the courts and advocating for continued proactivity on the part of the judiciary in the absence of any statutory or corporate self-governance developments for tackling supply chain human rights abuses.

Globalisation and Supply Chain Human Rights Abuses

Within the corridors of the World Trade Organisation and the offices of the world's largest companies a fear during recent decades has been that '*geographic fractionalisation [threatened] the multilateral trade arena*.'[9] To combat this fractionalisation the global economy has come to be dominated by systems of organisation which seek to reduce inconvenient sources of trade friction and to export production to countries and regions where labour and production costs can be reduced as far as possible. [10] Transnational companies have accrued vast capital and political might in this process, leading them to exert pressure on 'less-developed' countries to create attractive regulatory and legal frameworks for investment.[11] In this context, extractive[12] and manufacturing[13] industries- with well documented risks of human

Boersma 'The Governance of Global Value Chains: Unresolved Human Rights Environmental and Ethical Dilemmas in the Apple Supply Chain' [2015] 143 Journal of Business Ethics 111.

[7] *Vedanta Resources Plc v Lungowe* [2019] UKSC 20.

[8] *Okpabi v Royal Dutch Shell Plc (RDS)* [2021] UKSC 3.

[9] Donatella Alessandrini, 'The Time that Binds the 'Trade-Development' Nexus in International Economic Law' [2020] 12 Trade, Law and Development, 625; Daniel Kinderman, 'Free Us Up So We Can Be Responsible!' The Co-Evolution of Corporate Social Responsibility and Neo-Liberalism in the UK 1977-2010' [2012] 10 Socio-Economic Review, 29.

[10] Gary Gereffi, 'Global Value Chains and International Competition' [2011] 56 The Antitrust Bulletin, 36; Gary Gereffi, *Global Value Chains and Development: Redefining the Contours of 21st Century Capitalism* (Cambridge University Press 2018); Donatella Alessandrini, 'The Time that Binds the 'Trade-Development' Nexus in International Economic Law' [2020] 12 Trade, Law and Development, 625.

[11] Gwynne Skinner, 'Rethinking Limited Liability of Parent Corporations for Foreign Subsidiaries' Violations of International Human Rights Law' [2015] 72 Washington and Lee Law Review, 1769; Elisa Giuliani and Chiara Macchi, 'Multinational corporations' economic and human rights impacts on developing countries: a review and research agenda' [2014] 38 Cambridge Journal of Economics, 479.

[12] Penlope Simons and Audrey Macklin, *The Governance Gap: Extractive Industries, Human Rights and the Home State Advantage* (Routledge 2014); S James Anaya, 'Report of the special rapporteur on the rights of indigenous peoples on extractive industries and indigenous peoples' [2015] 32 Arizona Journal of International and Comparative Law, 110; Isabel Feichtner, Markus Krajewski and Ricarda Roesch, *Human Rights in the Extractive Industries* (Springer 2019).

[13] Elisa Giuliani and Chiara Macchi, 'Multinational corporations' economic and human rights impacts on developing countries: a review and research agenda' [2014] 38 Cambridge Journal of Economics, 479; Gay Seidman, *Beyond the Boycott: Labor Rights, Human Rights, and Transnational Activism* (Russel Sage Foundation 2007).

rights abuses- operate in legal systems which are unresponsive to victims, depriving them of redress.[14]

This process of human rights abuses becoming increasingly prevalent via the disaggregation of the global economy is multifaceted.[15] Human rights abuses are fostered by the systematic incentivisation of weak regulatory protections[16], the unbalancing effects of capital on legal systems in 'developing' countries[17] and the fluid environment of product sourcing occurring against a backdrop of governance and oversight gaps.[18] Concomitant with this process is the positive feedback loop to which victims of human rights abuses are subject. Already in precarious positions regarding their future employment and financial certainty, once subject to the effects of human rights abuses, victims find themselves devoid of any potential to litigate large organisations even should a viable regulatory framework exist for such action.[19] This inability is a product of the inhibitive disparity in available resources between the organisation perpetrating the human rights abuses and the victim. This is worsened by the unavailability of legal aid equivalents in most countries in which supply chain human rights abuses occur.[20] This has the effect of a frequently insurmountable barrier to justice even where a potential remedy should exist at law.

Acknowledgement of these issues has existed within the conceptualisations of international regulatory bodies, charities, governments, and business for decades. A litany of potential remedies has been mooted to redress the balance in favour of reconnecting the separation of supply-chain liability for human rights abuses from purchaser profit.[21] In a regulatory vacuum which covers issues of jurisdictional overlap and boundaries, one might suppose that legislative developments are warranted. However, no such statutory developments seem likely in coming years. The most recent Join Committee Special Report,[22] a valuable resource for analysis of governmental approaches to human rights abuses in the business sector with comprehensive deconstructions of the National Action Plan, makes no references to a need for new statutory provisions. The key source of hope for governmental action was a recommendation to fast-track the Modern Slavery (Transparency in Supply Chains) Bill.[23] However, governmental responses indicated scepticism of any new primary legislation in the

[14] Thomas Clarke and Martijn Boersma, 'The Governance of Global Value Chains: Unresolved Human Rights, Environmental and Ethical Dilemmas in the Apple Supply Chain' [2015] 143 Journal of Business Ethics, 111.
[15] ibid.
[16] Donatella Alessandrini, 'The Time that Binds the 'Trade-Development' Nexus in International Economic Law' [2020] 12 Trade, Law and Development, 625.
[17] Juan José Álvarez Rubio and Yiannibas Katerina, *Human Rights in Business; Removal of Barriers to Access to Justice in the European Union* (Taylor & Francis 2017), 41.
[18] Thomas Clarke, and Martijn Boersma, 'The Governance of Global Value Chains: Unresolved Human Rights, Environmental and Ethical Dilemmas in the Apple Supply Chain' [2015] 143 Journal of Business Ethics, 111.
[19] Alex Newton, *The Business of Human Rights: Best Practice and the UN Guiding Principles* (Routledge 2019) 5.
[20] Sheldon L Leader and Anil Yilmaz-Vastardis, 'Improving Paths to Business Accountability for Human Rights Abuses: A Legal Guide' [2017] Essex Business and Human Rights Project 6 <https://www1.essex.ac.uk/ebhr/documents/Improving-Paths-to-Accountability-for-Human%20Rights-Abuses-in-the-Global-Supply-chains-A-Legal-Guide.pdf> accessed 22 October 2021.
[21] Pierre Schmitt, *Access to Justice and International Organizations: The Case of Individual Victims of Human Rights Violations* (Edward Elgar Publishing Limited 2017).
[22] Joint Committee on Human Rights, *Human Rights and Business 2017: Promoting Responsibility and ensuring accountability* (2016-17, HL 153, HC 443).
[23] Joint Committee on Human Rights (n 23), 113.

area of business human rights abuses but committed to 'considering the Bill as it passes through parliament.'[24] Over four years later, a new version of the Bill is only in its first reading in the House of Lords[25] with the previous version having failed to progress through the Commons after being dropped by the sponsoring Conservative MP. [26] There have been no further attempts to extend statutory obligations for businesses regarding extraterritorial human rights and no reference made in the government's most recent round of consultations.[27] These sources, and the paucity of other indications, are demonstrative of the lack of governmental concern for further statutory developments.

Within the context of this void suggestions have been made for case law extensions which hold the potential to pierce the regulatory schism of jurisdictional boundaries in the context of global supply chains. Amongst these suggestions have been the potential that 'soft law' obligations might form the basis for tortious duties of care, thus improving human rights protections. Foremost amongst these have been the United Nations Guiding Principles[28] (UNGP) and the obligations therein for states to protect human rights. It has been proposed that domestic law developments in signatory countries might allow for extraterritorial extensions of liability for parent companies.[29] The conceptual basis for such a development is a 'hardening' of the *'social licence to operate'* into a legal duty wherein UNGP due diligence obligations might form a ground for establishing liability for supply chain human rights abuses.[30] Alternatively, it has been suggested that developments of corporate self-regulatory codes of conduct- increasingly frequent, companywide procurement, human rights and environmental policies which cover a spectrum from marketing devices to stringent supply chain wide standards- might hold the potential for case law extensions of human rights protections.[31] To date, such developments have not been actionable at a scale capable of

[24] Joint Committee on Human Rights, *Human Rights and Business 2017: Promoting responsibility and ensuring accountability: Government Response to the Committee's Sixth Report of Session 2016–17* (2017-19, HC 686) 8.
[25] HL Deb 21 July 2017, vol 783.
[26] UK Parliament, 'Modern Slavery (Transparency in Supply Chains) Bill [HL]' (Parliamentary Bills, 30 Nov 2016) <https://bills.parliament.uk/bills/1779/news> accessed 29 October 2021
[27] Home Office, 'Transparency in supply chains consultation. Government Response' (Consultations, 22 Sep 2020) <https://www.gov.uk/government/consultations/transparency-in-supply-chains> accessed 29 October 2021.
[28] United Nations, 'Guiding Principles on Business and Human Rights, UN Doc. A/HRC/17/31,21 March 2011, as endorsed by the Human Rights Council in Resolution 17/4 of 16' (*United Nations*, June 2011) <https://www.ohchr.org/documents/publications/guidingprinciplesbusinesshr_en.pdf> accessed 22 October 2021.
[29]Nicholas Bueno, 'Corporate liability for violations of the human right to just conditions of work in extraterritorial operations' [2017] 21 The International Journal of Human Rights, 565
[30] Claire Bright and others, 'Toward a Corporate Duty for Lead Companies to Respect Human Rights in Their Global Value Chains' [2020] 22 Business and Politics, 667.
[31] Deborah Leipziger, *The Corporate Responsibility Code Book* (3rd edn, Routledge 2017); Archie B Carroll, 'The pyramid of corporate social responsibility: Toward the moral management of organizational stakeholders' [1991] 34 Business Horizons, 39; Archie B Carroll and Kareem M Shabana, 'The Business Case for Corporate Social Responsibility: A review of concepts, research and Practice' [2010] 12 International Journal of Management Reviews, 85; Archie B Carroll, 'Carroll's pyramid of CSR: taking another look' [2016] 1 International Journal of Corporate Social Responsibility, 3.

reducing human rights abuses in global supply chains nor have they been tested in the English and Welsh legal system. [32]

Background to *Okpabi* and *Vedanta*

It is against the backdrop of decades of relative ineffectiveness in developments of regulatory protections against supply chain human rights abuses that the Supreme Court cases of *Vedanta* and *Okpabi* threw open new possibilities for the victims of human rights abuses.

Vedanta concerned a claim on the part of 1826 Zambian citizens for repeated discharge of toxic matter into local watercourses over a period of 15 years.[33] The citizens pursued Vedanta, domiciled in the UK, as the 'anchor defendant' for the conduct of Kongola Copper Mines plc- itself-regarded subsidiary. [34] *Okpabi* was a case concerning some 42,500 citizens of Nigeria contending that Royal Dutch Shell should be found to be responsible for the actions of their subsidiary (The Shell Petroleum Company of Nigeria Limited) for oil leaks from pipelines and infrastructure which polluted tracts of land and water sources which supported the appellants.[35]

Vedanta at its root represented a challenge to whether the English courts might exercise jurisdiction over a foreign subsidiary of a domestic company.[36] However, in so doing, the Supreme Court necessarily had to determine whether the previous leading case (*Chandler*[37])in the finding of a duty of care on the part of a parent company for the conduct of a subsidiary- - remained the sole approach for determining a duty of care in parent-subsidiary dynamics.

The argument of the defendant pertaining to this issue of law was that *Chandler* represented an exceptional circumstance and a finding of a duty of care on the part of a parent company can only be identified in circumstances where a case's factual matrix is similarly exceptional.[38] The Supreme Court, in rejecting this argument, determined that 'no novel extension' of the law of tort was required in the identification of a duty of care. All that needed to be applied was the exercise of the general principles of the law of tort.[39] In making such a finding the Supreme Court was the first court globally to identify a possibility for a parent company to be able, in the correct circumstances, to possess a common law duty of care to the victims of the conduct of a subsidiary in an extraterritorial scenario of this kind.[40]*Okpabi*

[32] Sheldon L Leader and Anil Yilmaz-Vastardis, 'Improving Paths to Business Accountability for Human Rights Abuses: A Legal Guide' [2017] Essex Business and Human Rights Project <https://www1.essex.ac.uk/ebhr/documents/Improving-Paths-to-Accountability-for-Human%20Rights-Abuses-in-the-Global-Supply-chains-A-Legal-Guide.pdf> accessed 22 October 2021.
[33] *Vedanta* (n 8), [1].
[34] Samantha Hopkins, 'Vedanta Resources plc and Another v Lungowe and Others' [2019] 70(3) Northern Ireland Legal Quarterly 371, 371.
[35] *Okpabi* (n 9), [3-4].
[36] Lucas Roorda and Daniel Leader, 'Okpabi v Shell and Four Nigerian Farmers v Shell: Parent Company Liability Back in Court' [2021] 6 Business and Human Rights Journal 2, 368.
[37] *Chandler v Cape plc* [2012] EWCA Civ 525, [2012] 1 W.L.R. 3111.
[38] Lucas Roorda and Daniel Leader, 'Okpabi v Shell and Four Nigerian Farmers v Shell: Parent Company Liability Back in Court' [2021] 6 Business and Human Rights Journal 2, 368.
[39] *Vedanta* (n 8), [51].
[40] Robert McCorquodale, 'Parent Companies Can Have a Duty of Care for Environmental and Human Rights' (Cambridge Core Blog, 11 April 2019). <https://www.cambridge.org/core/blog/2019/04/11/parent-companies-can-have-a-duty-of-care-for-environmental-and-human-rights-impacts-vedanta-v-lungowe/> accessed 22 May 2021.

reaffirmed the reasoning in *Vedanta*, namely that no special category for tortious liability has been created and that in determining the existence of a duty of care, the proper approach is to apply established principles of common law.[41] In the instant case the appellants asserted that it was proper to determine such a duty due to the degree of control which Royal Dutch Shell exerted over their subsidiary, the implementation of a global policy which the subsidiary breached, the existence of officers of the parent company holding oversight responsibilities, and the establishment of a Committee for ensuring the compliance of the subsidiary regarding Environmental protections.[42] All these were deemed by the Supreme Court to be possible mechanisms for establishing the requisite proximity for determining a duty of care. [43]

Effects of *Vedanta* and *Okpabi* on Future Human Rights Litigation

Whilst both *Vedanta* and *Okpabi* pertained to actions concerning environmental damage they are replete with possibilities for developing a framework for human rights litigation that might go some way in redressing the aforementioned regulatory lacunae which have fostered human rights abuses in global supply chains. It has been identified that the processes for evidencing and determining a duty of care on the part of a parent company for the conduct of a subsidiary could be applied in finding a duty where a subsidiary has caused, through its negligence, human rights abuses to occur.[44] This represents a seismic shift in the jurisprudence of the English and Welsh legal system. Prior to *Vedanta* and *Okpabi*, the twin barriers of discrete jurisdictional boundaries and an inability to pierce the parent-subsidiary corporate veil, save in exceptional circumstances, precluded the bringing of action against a UK domiciled company for human rights abuses.[45] Having established and affirmed that no special process exists for identifying such a duty of care, nor should jurisdictional boundaries sublate any duty, the Supreme Court has thrown open the possibility for human rights claims in parent-subsidiary circumstances.

This is significant for a number of reasons. The robustness of the English and Welsh legal system with regards human rights abuses and the well-developed common law principles governing this area of law represent a regulatory framework for victims to employ which is substantially more rigorous than can be found in many countries at the production end of global supply chains. [46] This means that the avenues of redress and recompense which victims might access are greatly increased. In addition, the English and Welsh legal system contains various avenues for support of the victims which serve to balance the chasm in resources and legal expertise which might exist should such cases be brought in victims' parent jurisdictions. Examples of this include the potential to access legal aid and the presence of law firms with

[41] *Okpabi* (n 9), [27].
[42] *Okpabi* (n 9), [29].
[43] *Okpabi* (n 9), [27].
[44] Sheldon L Leader and Anil Yilmaz-Vastardis, 'Improving Paths to Business Accountability for Human Rights Abuses: A Legal Guide' [2017] Essex Business and Human Rights Project <https://www1.essex.ac.uk/ebhr/documents/Improving-Paths-to-Accountability-for-Human%20Rights-Abuses-in-the-Global-Supply-chains-A-Legal-Guide.pdf> accessed 22 October 2021.
[45] Richard Meeran, 'Multinational Human Rights Litigation in the UK: A Retrospective' [2021] 6(2) Business and Human Rights Journal 255.
[46] Samantha Hopkins and others, 'Okpabi and others v Royal Dutch Shell Plc and another [2021] UKSC 3; [2021] 72 Northern Ireland Legal Quarterly 1, 148.

the resources, expertise, and willingness to take on class action and individual cases of this nature against large multinational organisations.

The decisions also have the wider implication of demonstrating that the UK courts are willing to act proactively in seeking more global protections for human rights. The effects of this are twofold- by demonstrating a willingness to meet the UK's obligations under the UNGPs the courts have affirmed these obligations and have opened the door for regulatory and oversight obligations such as these to form the basis for future litigation.[47] Secondly, the courts have demonstrated to UK domiciled multinational corporations that technical legal barriers will not operate as an aegis behind which they might directly profit from, perpetrate, facilitate or allow human rights abuses to occur throughout the global supply chains which they control or from which they profit with sufficient proximity.[48] Exemplifying this breakdown of a technical barrier is the deterioration of the doctrine of *forum non conveniens* in its application by the courts of England and Wales in frustrating human rights claims.[49] This breakdown is predicted to serve as a force to discourage human rights abusive practises or oversight omissions on the part of multinationals, imposing liability for abuses on the organisations capable of implementing or engendering the changes necessary to vitiate the possibility for such abuses occurring in the future.[50]

Limitations of *Vedanta* and *Okpabi*

Whilst *Vedanta* and *Okpabi* are representative of a potentially significant development in efforts to tackle the human rights abuses perpetrated by UK domiciled multinationals in their supply chains, these two cases are by no means constitutive of a panacea. The effects they will have in severing the connection between globalised supply chains and the creation of circumstances productive of human rights abuses is likely to be limited by a number of factors.

A considerable issue in utilising the framework established in *Vedanta* and *Okpabi* in bringing human rights abuse claims is financial barriers to the viability of such claims. Whilst the provisions for victims to have access to mechanisms and provisions designed to improve their ability to bring claims and participate meaningfully in proceedings are more substantial in the English and Welsh legal system, the fundamental reality of claims of this nature is that they rely on the case being won such that the defendant pay the appellant's costs. This means that it is likely that solicitors will conduct a preliminary analysis of the likelihood of a claim's success which is more rigorous than for cases in other circumstances. This will logically eventuate certain claims which may have produced findings of human rights abuses not being brought.

Furthermore, claims will be reliant upon the finances of the defendant which is not always as robust as might be expected of a multinational corporation. An example thereof is

[47] Samantha Hopkins, 'Vedanta Resources plc and Another v Lungowe and Others' [2019] 70 Northern Ireland Legal Quarterly 3, 371.
[48] Richard Meeran, 'Multinational Human Rights Litigation in the UK: A Retrospective' [2021] 6 Business and Human Rights Journal 2, 255.
[49]Vlex 'Ngcobo v Thor Chemicals Holdings Ltd & Desmond Cowley' (Times L Rep 10 November 1995) <https://vlex.co.uk/vid/ngcobo-and-others-v-806088957> accessed 23 May 2021.
[50] Gwynne Skinner, 'Rethinking Limited Liability of Parent Corporations for Foreign Subsidiaries' Violations of International Human Rights Law [2015] 72 Washington and Lee Law Review, 1769.

Cape plc's financial state which resulted in a reduced settlement in *Chandler v Cape plc*.[51] Moreover, efforts can be made by multinational corporations to frustrate efforts to recoup damages by attempting to obviate tracing efforts against funds; employing the vagaries of corporate structures to hide funds offshore. An example thereof is the efforts by Thor Chemicals to shield their liquidity via holding companies following a damage finding arising out of *Ngcobo and Others v Thor Chemicals Holdings Ltd and Others.* [52] Finally, the complexity of the proceedings can mean that the costs incurred are of greater risk of being deemed unrecoverable.[53] This final limitation should be subject to the corollary that the principle of proportionality somewhat redresses this risk by application of relevant determining factors including whether the complexity of the litigation and the wider public importance of the proceedings should increase the recoverable costs.[54]

The principles which underpin the finding of a duty of care are necessarily limited at common law for that overarching fear of opening the proverbial floodgates. To this end, a *Vedanta* and *Okpabi* framework for liability is limited by the requirements for determining the requisite proximity for the establishment of a duty of care. It is unavoidable in this context that certain modes of production which are highly causative of human rights abuses will fall without the bounds of determining such a duty of care. Examples of such modes include the increasingly stratified and offshored purchasing processes which characterise the manufacturing industry in particular.[55] In such systems, a UK domiciled corporation will contract with a series of outsourcing organisations which produce sufficient remoteness of relationship that the finding of a duty of care becomes almost completely vitiated.[56] In primary extractivist industries, as considered in the *Vedanta* and *Okpabi* factual profiles, the potential for ephemeral and shifting modes of supply chain organisation are obviated by the need to maintain a consistent presence at a given locale for the extraction of the resource in question. In the context of an increasingly short term and multi-layered approach to production, manufacturing, and procurement, there remains the possibility that the relationships therein are insufficient for the creation of duties of care. It is possible that the *Vedanta* and *Okpabi* decisions alone, without the addition of new regulatory obligations to govern supply chain relationships or to impose oversight obligations on the part of organisations at the head of those chains, will be incapable of precluding human rights abuses in those supply chains.[57]

Analysis of the Future of *Vedanta* and *Okpabi* Claims

[51] *Chandler v Cape plc* [2012] EWCA Civ 525, [2012] 1 W.L.R. 3111.
[52] ECA 10 Nov 1995 Times L Rep 10 November.
[53] Richard Meeran, 'Multinational Human Rights Litigation in the UK: A Retrospective' [2021] 6 Business and Human Rights Journal 2, 255.
[54] ibid.
[55] Mark A Geistfeld, 'The Law and Economics of Tort Liability for Human Rights Violations in Global Supply Chains' [2019] 10 Journal of European Tort Law 2, 130.
[56] Simons P and Macklin A, *The Governance Gap: Extractive Industries, Human Rights and the Home State Advantage* (Routledge 2014); Giuliani E and Macchi C, 'Multinational corporations' economic and human rights impacts on developing countries: a review and research agenda' [2014] 38 Cambridge Journal of Economics, 479.
[57] Carolijn Terwindt and others, 'Supply chain liability: pushing the boundaries of the common law?' [2018] 8 Journal of European Tort Law, 261.

It remains to be seen to what extent *Vedanta* and *Okpabi* will produce a new wave in human rights litigation and help to redress the organisational effects of the global network of supply chains with regards to said abuses. A possibility which has been suggested in the aftermath of *Vedanta* and *Okpabi* is that extensions of duties of care will simply serve to disincentive any efforts at corporate oversight in the supply chain for fear of demonstrating the control or proximity requisite for establishing a duty of care.[58] However, this analysis is insular to a more significant counter directional movement in the global economy. Increasingly, UK domiciled organisations are under pressure to demonstrate their corporate responsibility via the introductions of corporate codes of conduct.[59]

With regards these optional social responsibility commitments it is becoming apparent that, regardless of any conceptualisation on the part of the engendering organisation that these are merely 'corporate puff', legal authorities are coming to consider these as creating real legal responsibilities.[60] There are also increasing instances of obligatory corporate due diligence commitments perhaps best exemplified by the due diligence obligations under the UNGPs.[61] The direction of movement therefore is one towards greater commitment to oversight in the supply chain rather than in the direction of greater distancing. Against the backdrop of these commitments and obligations multinational corporations are increasingly less able to distance themselves from their supply chains to escape tortious duties of care which might otherwise have been identified.[62]

Whilst the thrust of this article has been to demonstrate the significance of the changes instigated by *Vedanta* and *Okpabi* and the potential that this must reduce the instances of human rights abuses in supply chains and to provide recourse for victims, it is worth examining a further militating argument in favour of future judicial implementations of the reasoning employed by the Supreme Court. In their consideration of the importance of the public interest in any extension of tortious liability it is worth denoting that extended liability for human rights abuses can reward socially responsible companies and incentivise sustainable business practises.[63] Appetite exists for such expansions, evidenced by analysis of businesses

[58] ibid.
[59] Deborah Leipziger, *The Corporate Responsibility Code Book* (3rd edn, Routledge 2017); Archie B Carroll, 'The pyramid of corporate social responsibility: Toward the moral management of organizational stakeholders [1991] 34 Business Horizons, 39; Archie B Carroll and Kareem M Shabana, 'The Business Case for Corporate Social Responsibility: A review of concepts, research and Practice' [2010] 12 International Journal of Management Reviews, 85; Archie B Carroll, 'Carroll's pyramid of CSR: taking another look' [2016] 1 International Journal of Corporate Social Responsibility, 3.
[60] Claire Bright and others, 'Toward a Corporate Duty for Lead Companies to Respect Human Rights in Their Global Value Chains' [2020] 22 Business and Politics, 667.
[61] United Nations, 'Guiding Principles on Human Rights' [2011] <https://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf> accessed 22 May 2021; John Sherman III, 'Human Rights Due Diligence and Corporate Governance' Corporate Responsibility Initiative' (Harvard Kennedy School, Working Paper No.79 June 2021) <https://www.hks.harvard.edu/sites/default/files/centers/mrcbg/files/CRI_WP_79_Final.pdf> accessed 30 October 2021.
[62] Sheldon L Leader and Anil Yilmaz-Vastardis, 'Improving Paths to Business Accountability for Human Rights Abuses: A Legal Guide' (Essex Business and Human Rights Project, 2017) <https://www1.essex.ac.uk/ebhr/documents/Improving-Paths-to-Accountability-for-Human%20Rights-Abuses-in-the-Global-Supply-chains-A-Legal-Guide.pdf> accessed 22 October 2021.
[63] Carolijn Terwindt and others, 'Supply chain liability: pushing the boundaries of the common law?' [2018] 8 Journal of European Tort Law 261,12.

themselves who value the legal certainty that would be afforded by a defined position as opposed to the uncertainty contemporarily extant. [64] Attempts to frame developments as anti-business or unleashing potentially unlimited liability are fundamentally unfounded in the context of the careful and incremental deliberations exemplified in *Vedanta* and *Okpabi* and the lengths that the courts went to ensure that they were not creating a new category of liability.

Conclusion

Vedanta and *Okpabi* are indicative of shifting judicial attitudes towards extra-territorial tortious liability for environmental and human rights abuses. They are also representative of one of the most significant global transitions in approach. Considering the recency of the decisions, it remains to be seen how widely the approach of the Supreme Court will be followed and whether the barriers to victims pursuing human rights claims deter more widespread efforts to utilise the framework presented by the two cases. Furthermore, the cases are not an overarching cure to supply chain human rights abuses, replete as they are with conceptual limitations which are not capable of filling the regulatory void on their own. Nonetheless, the potential of *Vedanta* and *Okpabi* should not be too maligned.

Operating within common law incrementalism and conceptual limitations they are likely to prove of considerable utility. Future possibilities for symbiotic litigation approaches, utilising soft law or self-governance codes and commitments to extend the ability to determine duties of care, mean that the full extent of possibilities presented by the two cases has yet to be fully explored. The Supreme Court has not filled the regulatory void but victims of human rights abuses and those seeking a more responsible global system of supply should be heartened by these developments, their future potential, and the growing appetite for change that they represent.

[64] Lisa Smit and others, 'Business Views on Mandatory Human Rights Due Diligence Regulation: A Comparative Analysis of Two Recent Studies' [2020] 5 Business and Human Rights Journal, 261.

On Thin Ice – Clinical Turned Criminal: Reforming the law on Medical Manslaughter

Lara Ozyer

Abstract

The laws surrounding medical manslaughter construct a confusing picture of accountability for professionals. Thus, many academics and prosecutors have deemed it a failure for its uncertainty and inability to differentiate between inadvertence and intentional disregard. There is unease regarding the readiness of the current law to assign criminal liability to professionals and a failure to recognise the high-pressure environment in which sub-standard practice occurs.

This paper explores the shortcomings of the current legal framework and the detrimental impact on healthcare professionals who get caught up within this unrefined area of law. A culpability-based reform will be proposed to shift the threshold for criminal liability from negligence to recklessness. This ensures that, rather than focusing on death to be inflicted, courts evaluate all case factors, including the defendant's state of mind and intentions surrounding their conduct. The difficulties raised by the proposed reform will also be addressed. Problems with defining and proving subjective recklessness, doctors' incompetence in eliminating foreseen risks, and the lack of deterrence effect recklessness create compared to gross negligence are among them. Ultimately, it will be argued that subjective reckless manslaughter is the appropriate place to set the bar for liability. A reform of this nature will undoubtedly contribute to transforming the law into the high-performing legal system it aspires to be.

Introduction

The criminal doctor makes a good story.[1] The convictions of doctors *Sellu*[2] and *Bawa-Garba*[3] revived the debate about whether doctors should be subject to criminal law and how they should be punished for gross negligence manslaughter. In response to concerns that prosecutions take place 'at the least excuse', Dr Michael Powers claimed that the criminal law should have a role in medical manslaughter. He states that 'whether a doctor is going to find himself on a manslaughter charge will depend on how much thought…care, and… attention he is giving to what he does....'[4] Suppose Dr Powers' contention is accurate, and poor clinical treatment should be prosecuted; in that case, it is vital to address what conduct should be pursued under criminal law and the level of culpability that warrants a criminal sanction.

[1] Margot Brazier and others, 'Improving Healthcare Through the Use Of 'Medical Manslaughter'? Facts, Fears and The Future' [2016] 22 Clinical Risk, 88.
[2] *R v Sellu* [2016] EWCA Crim 1716.
[3] *R v Bawa-Garba* [2016] EWCA Crim 1841.
[4] Dr Michael J Powers, 'Manslaughter - How Did We Get Here?' [2005] 73 Medico-Legal Journal, 123.

Concerns of the current law on gross negligence manslaughter originate from its arbitrary nature, which renders even the rare prosecutions futile as professionals are unaware of what they are doing so wrong that they could be held criminally accountable. Their complaints can be broken into two categories:

(1) The scope of the offence is uncertain as it does not identify the 'bad' doctors. Instead, it unearths negligent doctors whose patients are unfortunate enough to die.

(2) The culpability element of this offence is absent, resulting in the prosecution and criminalisation of doctors who act in good faith.

This paper examines the current law on medical manslaughter, highlighting its inadequacies and arguing that it is insufficient and overly burdensome on medical professionals. The inability of the law to establish a clear-cut definition and distinguish between unintentional errors and blatant negligence deems unjust outcomes inevitable. A culpability-based reform that establishes recklessness as the benchmark for criminal liability will be explored. It will acknowledge that although such reform is desirable in that it seeks to impose legal certainty by clarifying the ambiguous Adomako test and the Crown Prosecution Service's (CPS) vague approach to medical manslaughter cases, there are concerns associated with it which will be addressed in turn. Ultimately, it will be concluded that subjective reckless manslaughter is the appropriate place to set the bar for liability.

Brief Context: The Current Law on Medical Manslaughter

The current English law on medical manslaughter seeks to punish negligent acts of medical professionals that result in death. The common law offence of gross negligence manslaughter is broadly defined as 'where death is a result of a grossly negligent act or omission on the part of the defendant.'[5] A manslaughter conviction may be justified when a healthcare professional did not intend to kill or cause grievous bodily harm to a patient, but their extreme carelessness resulted in a patient's death while acting within the scope of their duty of care.

The House of Lords confirmed the legal test for gross negligence manslaughter in *Adomako*.[6] It is a four-staged test in which all the essential elements must be established to warrant a criminal conviction. These are as follows:

(1) The defendant must owe a duty of care towards the deceased.

(2) The defendant must have breached that duty of care.

(3) The breach must have caused or significantly contributed to the death.

(4) The breach must be characterised as gross negligence and therefore considered a crime.[7]

In such cases, the jury is directed to Lord Mackay's speech in *Adomako*, which established the legal principle that 'the defendant's conduct must have departed from the proper standard of care incumbent upon him.'[8] It should be noted that where an individual believes they acquire a

[5] The Crown Prosecution Service, 'Gross Negligence Manslaughter' (Cps.gov.uk, 2019) <www.cps.gov.uk/legal-guidance/gross-negligence-manslaughter> accessed 5 January 2022.
[6] *R v Adomako* [1995] 1 AC 171.
[7] *R* v *Adomako* [1995] 1 AC 171 at 187.
[8] ibid.

particular skill or knowledge, their conduct must be reviewed against the reasonably competent professional within that field of expertise, as is the case for healthcare professionals.[9] Although these issues will be discussed further, worthy of brief note is that this principle does not acknowledge the administrative discrepancies that medical professionals may encounter during healthcare delivery. While the desire to criminalise dangerous conduct is understandable, the inevitable nature of medical errors is overlooked. This is especially detrimental to doctors who act in good faith yet make momentary mistakes under pressure. This lends credence to the argument that recklessness is superior to negligence. It strikes a delicate balance between identifying negligent doctors and allowing leeway for innocent doctors; the law must not exacerbate the problem.

Issues with the Current Law on Medical Manslaughter

Mistake or negligence?

Predominantly, a significant criticism in academic reviews levied towards medical manslaughter is the law's emphasis on harm rather than moral culpability.[10] It is argued that there is a lack of a clear and precise definition of the law on gross negligence manslaughter. An article published by the Bar Council asserts that 'the law fails to make the critical distinction between flagrant negligence and fleeting mistake.'[11] The inability of the law to distinguish between incompetent doctors who make continuous poor decisions and those who make unintentional momentary errors under stress highlights the persistent issues of uncertainty and inconsistency. This sets a troubling and unpredictable precedent for defendants; as Andrew Ashworth states, 'people must be able to find out what the law is and factor it into their practical decisions.'[12]

The case of *Prentice* and *Sullman*[13], whose appeals were heard in *Adomako*, is a remarkable example of good doctors who have become entangled in this area's imprecise law. This case concerned two untrained doctors tasked with administering chemotherapy to a young boy. Due to their inexperience and inadequate monitoring, the wrong medicine was administered, resulting in the boy's death. Consequently, both doctors were convicted of manslaughter as it was irrelevant in the eyes of the law that they were 'far from being bad

[9] Emily Finch and Stefan Fafinski, *Criminal Law* (6th edn, Pearson, 2016).
[10] Andrew Sanders and Danielle Griffiths, *Medicine, Crime, and Society* (1st edn, Cambridge University Press, 2013), 117-158; Oliver Quick, 'Medicine, mistakes and manslaughter: A criminal combination?' [2010] 69 The Cambridge Law Journal; Michelle Robson, Jon Maskill and Warren Brookbanks, 'Doctors Are Aggrieved—Should They Be? Gross Negligence Manslaughter and The Culpable Doctor' [2020] 84 The Journal of Criminal Law; Bar Council, 'When Clinical becomes Criminal: Reforming Medical Manslaughter', (The Bar Council) <https://www.barcouncil.org.uk/media/627460/> accessed 10 December 2021.
[11] Bar Council, 'When Clinical becomes Criminal: Reforming Medical Manslaughter', (The Bar Council) <https://www.barcouncil.org.uk/media/627460/> accessed 10 December 2021.
[12] Andrew Ashworth, 'Ignorance of the Criminal Law, and Duties to Avoid it', [2011] The Modern Law Review, 1.
[13] *R v Prentice, Sulman, Adamako, Holloway* [1993] 4 All ER 935.

men.'[14] Similarly, Dr Ubani[15] flew to England under 'tremendous stress and exhaustion' to work as an after-hours doctor. He accidentally killed his patient due to improperly administering a drug he had not acquired adequate general training or knowledge. Subsequently, he was convicted in his home country Germany of killing his patient in England. Both cases mentioned resulted from system failures and insufficient training for administering specific medications. However, despite lacking the malicious intent to cause death, they were prosecuted and labelled criminals. Indeed, both cases involved some level of moral culpability, as essentially, their actions resulted in death, deeming them blameworthy. However, whether their level of guilt is deserving of a criminal sanction is questionable, especially in the light of examples of flagrant negligence, such as *Adomako*, in which a doctor failed to notice a dislodged oxygen tube during surgery for more than five minutes.

Professionals acting in good faith should not be labelled criminals.[16] However, the element of intentionally making a wrong choice, the prerequisite mens rea of all serious crimes, is absent in gross negligence manslaughter.[17] A doctor aware that they are taking an unnecessary risk is reckless.[18] This is the most culpable state of mind associated with gross negligence manslaughter, short of outright intent to cause harm. However, determining whether a doctor is deserving of a criminal conviction for making an inadvertent mistake that results in the death of a patient is not as straightforward. Critics argue that to use negligence alone, no matter how gross, to convict medical practitioners of a criminal offence is unfair; there must also be demonstrable culpability.[19]

While it is acceptable to find doctors negligent if they cause harm to a patient by failing to meet professional standards, it is less palatable to find them criminally liable if their intentions were not additionally culpable. As such, the current offence does not identify bad doctors but merely unearths those doctors whose patients are unfortunate enough to die. This sheds light on the shortcomings of the current law; it is improperly defined as it fails to target those doctors who act negligently, and society should criminalise. This gives weight to the argument that recklessness is the appropriate place to set the bar for liability because it ensures that the defendant's state of mind is taken into consideration, pinpointing only those doctors deserving of a criminal sanction.

Quick notes that even for the CPS, maintaining the framework of medical manslaughter is difficult.[20] There is some trepidation regarding applying the full force of prosecution against doctors whose errors have had disastrous repercussions. This scepticism, even among prosecutors, demonstrates that the law has fundamental flaws that are overly harsh on

[14] Margaret Brazier and Amel Alghrani 'Fatal Medical Malpractice and Criminal Liability' [2009] 25 Journal of Professional Negligence 51, 56.
[15] 'Doctor Daniel Ubani Unlawfully Killed Overdose Patient' (*the Guardian*, 2010) <https://www.theguardian.com/society/2010/feb/04/doctor-daniel-ubani-unlawfully-killed-patient> accessed 10 January 2022.
[16] Michelle Robson, Jon Maskill and Warren Brookbanks, 'Doctors Are Aggrieved—Should They Be? Gross Negligence Manslaughter and The Culpable Doctor' [2020] 84 The Journal of Criminal Law.
[17] *R v Adomako* [2000] QB 796, [809].
[18] ibid., [15].
[19] ibid., [16].
[20] Oliver Quick, 'Prosecuting 'Gross' Medical Negligence: Manslaughter, Discretion, and the Crown Prosecution Service' [2006] Journal of Law and Society 33, 441.

professionals who operate in demanding and precarious circumstances. He also found a greater emphasis on the part of the CPS in prosecuting cases involving recurring failures and decisions to dismiss warnings.[21] Still, some examples of momentary errors have been pursued. Although prosecutions with the concept of culpability are the logical approach for the CPS, such an apparent pattern in prosecutions is insufficient as it undermines transparency. This further supports the notion that reckless manslaughter would be better than gross negligence. Amending the law to reflect our moral sense of culpability and punishing only those doctors who are genuinely guilty of a crime would result in a far more just outcome, hence, better attaining the purpose of the law on medical manslaughter.

Uncertainty

Uncertainty is another prominent issue entailed within the current law of medical manslaughter.[22] The test for gross negligence manslaughter and the CPS's approach to prosecution lacks clarity and certainty, casting doubt on their compatibility with the rule of law. The test for distinguishing between mere negligence and gross negligence was established in *Adomako*, where it was acknowledged that there must be a breach of duty that causes death. According to Lord Mackay, based on the extent to which the defendant's conduct deviated from the standard of care, 'the jury must consider whether that breach of duty should be characterised as gross negligence and therefore as a crime.'[23] However, apart from Lord Mackay's observation that the conduct must be 'so bad' as to constitute a criminal act, limited guidance is provided on the elusive notion of 'grossness' entails.[24]

According to a report produced by the Law Commission highlighting the faults of the current law on manslaughter, indifference and gross negligence are defined as virtually the same offence. They propose the necessity of a 'clear and robust difference between offences of varying degrees of gravity' to eliminate this ambiguity. Although the essential elements of gross negligence manslaughter are clear, the circularity created by a definition of gross negligence being criminal had previously been outlined as a concern by Lord Mackay in *Adomako*.[25] Judge LJ acknowledged in *Misra*: 'it is true that to a certain extent this involves an element of circularity… However, the essence of the matter, which is supremely a jury question, is whether, having regard to the risk of death involved, the conduct of the defendant was so bad in all the circumstances as to amount in their judgement to a criminal act or omission…..'[26] As such, this 'badness' should be a distinct and comprehensible term since it is the main component of the offence that converts mere negligence into criminal conduct.

Despite this, the judiciary has overlooked opportunities to amend and clarify this test, providing validity to Griffith and Sanders' assertion that the law on medical manslaughter is

[21] ibid.
[22] Melinee Kazarian, Danielle Griffiths, and Margaret Brazier, 'Criminal responsibility for medical malpractice in France' [2011] Journal of Professional Negligence 27, 188.
[23] *R v Adomako* [1995] 1 AC 171, [25] (Lord Mackay).
[24] ibid [20].
[25] *R v Adomako* [1995] 1 AC 171, [7] (Lord Mackay).
[26] *R v Adomako* [1995] 1 AC 171, [187]; *Rowley* v *DPP* [2003] EWHC Admin 693, (Kennedy LJ).

'curiously indispensable.'[27] For example, in *Misra*,[28] the Court of Appeal upheld the *Adomako* test for gross negligence, following an appeal by two senior house officials against their manslaughter convictions, claiming their right to punishment without law under article 7 of the European Convention of Human Rights (ECHR) was violated. Essentially, under this article, a defendant cannot be convicted if the offence is not sufficiently defined for them to have had prior knowledge of it. However, as this 'certainty' does not have to be absolute, Judge LJ concludes that the law as it currently stands provides sufficient clarity and affirms that the test requires the jury to consider grossness and criminality as a single element as opposed to two separate questions.

Yet, the definition remains vague and leaves too much space for jury interpretation. In effect, the jury is left to decide on a legal question: when is negligence sufficiently gross to adjudicate criminality? The role of the judge is to direct the jury on the point of law, while the jury's function is to debate problems of fact, not law. If, as Lodge points out, 'following the imposition of a gross negligence manslaughter charge, the jury is tasked with filtering out those failures deserving of criminal sanction',[29] any degree of certainty is exceedingly improbable when it is dependent on the subjective opinion of a jury. This approach was dismissed by Judge LJ, who stated, 'the question for the jury is not whether the defendant's negligence was gross, and whether, additionally, it was a crime, but whether his behaviour was grossly negligent and consequently criminal.'[30] As Ashworth pointed out, 'is a distinction without a difference… it should not be the last word on the subject.'[31] Thus, implementing the *Adomako* test promotes inconsistency and does little to attain justice.

Furthermore, Quick highlights the problems that arose when prosecutors attempted to define gross negligence, which they could not do without referring to 'badness'[32]. Moreover, there appears to be no indication of a solid policy within this area of law, with prosecutors emphasising the need for the expertise of prosecutors and 'gut instinct.' Prosecutors tended to be highly reliant on their moral frameworks when prosecuting practitioners for whom there was an absence of recorded comparative information regarding their conduct. Although Hawkins assumes that all prosecutorial discretion is based on an individual's moral framework,[33] it is argued here that a clearly defined offence would reduce arbitrary decision making at the prosecution stage, encouraging transparency, and attaining justice more effectively. This gives weight to the reformulated test based on recklessness suggested in this paper; it would offer greater certainty at the prosecution stage by setting the liability bar at a precise level aiding the differentiation between varying degrees of gravity.

Proposed Reform: Shifting Negligence to Recklessness

[27] Danielle Griffiths and Andrew Sanders, 'The road to the dock: prosecution decision-making in medical manslaughter cases' [2013] Bioethics, medicine, and the criminal law 2, 117.
[28] *R v Misra & Srivastava* [2004] EWCA Crim 2375.
[29] Anne Lodge, 'Gross Negligence Manslaughter on The Cusp' [2017] 81 The Journal of Criminal Law.
[30] *R v Misra & Srivastava* [2004] EWCA Crim 2375, [62] (Judge LJ).
[31] Andrew Ashworth, *Principles of Criminal Law* (7th edn, Oxford University Press 2013), 293.
[32] ibid [18].
[33] Keith Hawkins, *Law as Last Resort: Prosecution Decision-Making in a Regulatory Agency* (OUP 2010).

Several reforms have been proposed in recent years, including the Law Commission's proposal to replace killing with gross carelessness.[34] Quick dismissed this as 'nothing more than a linguistic modernisation of the status quo,' which he claims fails to 'address the fundamental objection to negligence-based criminal liability.'[35] Instead, he favours Tadros' test, which focuses on the perpetrator's breach of duty, either by failing to research dangers or by being intentionally oblivious to the presence of a risk.[36] However, this approach could lead to the incarceration of doctors like Prentice and Sullman. They failed to examine the risks of injecting medications intrathecally, even though their fatal omission was due to ignorance, not indifference. Objectively, in contrast to the hypothetical 'reasonable person', they were reckless, but subjectively, given their inexperience and the circumstances they found themselves in, they were merely negligent.[37] Mullock takes a more satisfactory stance as he believes that subjective recklessness is the appropriate basis for liability.[38] According to McCall Smith, prosecutions should only be brought when reckless conduct and the perpetrator 'deliberately and culpably took a risk with their patients.'[39] Brazier states that, 'only such conduct pursued with disregard for the lives of others should merit punishment.'[40] Crosby has advocated for the replacement of gross negligence manslaughter with reckless manslaughter, based on a capacity-based approach that emphasises the need to 'appreciate a risk and the context in which the proscribed conduct occurred.'[41]

Thus, a reformulated test based on culpable recklessness, rather than a modified test for gross negligence, would operate better as negligence is an unsuitable basis for medical manslaughter.[42] Consideration of the professional's subjective mind and intentions is necessary for pinpointing and prosecuting only those doctors who have shown deliberate disregard for their patient's life. As stated at the outset of this paper, the existing negligence-based criteria fails to identify 'bad' doctors, whereas requiring this criterion would undoubtedly rule out doctors who were not negligent and hence do not deserve to be labelled criminals.

Brazier and Algharni's work in redefining gross negligence can be used to construct a new test.[43] Based on the model proposed, the amendments to the medical manslaughter test are inclusive of:

(1) Whether the doctor's conduct fell short of what would be considered responsible professional practice resulting in tort liability

[34] Law Commission, *Involuntary Manslaughter* (Law Com No 237, 1996), 22-24.
[35] Charles A Erin and Suzanne Ost, *The Criminal Justice System and Health Care* (OUP 2007), 47.
[36] Victor Tadros, 'Recklessness and the Duty to Take Care' in Stephen Shute and Andrew Simester 'Criminal Law Theory: Doctrines of the General Part' (Oxford University Press, Oxford 2002) 227–258.
[37] Cath Crosby, 'Gross Negligence Manslaughter Revisited: Time for a Change of Direction?' [2020] 84 Journal of Criminal Law 3, 288.
[38] Alexandra Mullock, 'Gross Negligence (Medical) Manslaughter and the Puzzling Implications of Negligent Ignorance: *Rose v R* [2017] EWCA Crim 1168' [2018] Medical Law Review, 346.
[39] ibid., [38], 336-349.
[40] Margaret Brazier and Neil Allen, 'Criminalising Medical Malpractice' in Erin and Ost (eds), The Criminal Justice System and Health Care (Oxford University Press, 2007), 27.
[41] ibid [34], 245.
[42] Alexander, Ferzan and Morse, *Crime and Culpability: A Theory of Criminal Law* (Cambridge University Press, 2009).
[43] Margaret Brazier and Amel Alghrani, 'Fatal Medical Error and Criminal Liability' [2009] 25 *Professional Negligence* 49.

(2) Whether the doctor was oblivious to the evident risk of significant injury to his patient (if yes, he has acted recklessly as he has failed his responsibility to his patient and is culpable for the harm caused)

(3) Whether the doctor was aware of such a risk yet chose to subject the patient to it even though no recognised medical benefit existed (unless there is substantial evidence of important mitigating factors, he has acted recklessly)

It should be noted that mitigation should be granted in cases involving incapacity and a lack of expertise to minimise the current law's excessive brutality on inexperienced and overworked doctors. However, if the doctor were aware that his inexperience or incapacity was likely to cause harm and still went ahead regardless, he would still be held culpably reckless. Therefore, the mitigation should consider all other external factors, and it is not an absolute defence to overly poor practice. This parallels the *Adomako* test, which refers to 'all the circumstances', not only the outcome of the conduct, and is vital given the law's contentious nature and its wide-ranging ramifications for doctors.

As such, this paper proposes that shifting the threshold of culpability within the law of medical manslaughter from negligence to recklessness eliminates the issues posed by the current legal framework. This would be a reasonable and consistent departure from the current law, which is uncontentious and well understood despite a lack of recent common law authority. To prevent law-making for the sake of law-making and uphold the notion of legal minimalism, recklessness should be integrated into the current law as seamlessly as possible. This would not be difficult as although the higher courts have shown preference to gross negligence; the term 'reckless' has always remained within the peripheral scope of gross negligence manslaughter. For instance, as suggested by Lord Mackay's advice in *Adomako*, there is an apparent inclination to conflate negligence with recklessness, as he stated that 'it is perfectly open to the trial judge to use the word 'reckless' in its ordinary meaning as part of his exposition of the law if he deems it appropriate.'[44]

This was supported in *Misra* where despite the ostensibly objective test (gross negligence), the Court of Appeal stated that evidence of the defendant's state of mind is irrelevant to the issue of gross negligence.[45] It was further asserted that often, there will be a 'critical factor in the decision', thereby upholding subjectivity.[46] As Quick suggests, it is not a coincidence that recklessness lurks in the background, possibly reflecting judicial reluctance to leave culpability at a level lower than gross negligence.[47] Due to their unease with the unfair and ambiguous objective test, prosecutors frequently look for subjective fault as a basis for their judgements of grossness. It is apparent that the judicial system desires to 'rediscover recklessness.'[48]

Justifying the Proposed Reform: A Better Alternative?

[44] *R v Adomako* [1995] 1 AC 171, [9] (Lord Mackay).
[45] *Attorney-General's Reference (No. 2 of 1999)* [2000] 3 All ER 182, (Rose LJ).
[46] ibid.
[47] Oliver Quick, Medicine, mistakes and manslaughter: A criminal combination?' [2010] 69 The Cambridge Law Journal.
[48] ibid., [168, 188].

Differing Gravities

Arguably, the current framework contradicts the rule of law. The test for gross negligence does not reveal to medical practitioners what the law expects of them and what constitutes a crime. The issues that prevail are the uncertainty and inconsistency that is failing the justice system and sets a concerning unpredictable precedent for defendants as they cannot be sure what constitutes criminal conduct that is punishable. Thus, the law is inefficient, as the true purpose of medical manslaughter law cannot be attained. This is because it is not only the bad doctors who are punished but also anyone who makes a mistake, which is inevitable in the healthcare sector. Hence, the law on gross negligence manslaughter has been criticised for punishing doctors irrationally and unjustly.

While some prosecutors argue it is wrong to prosecute doctors merely because their profession is inherently risky,[49] many academics have advocated that the offence be abolished entirely. However, allowing doctors to inflict death without facing any repercussions would be a step too far and undoubtedly undermine the purpose of criminal law as a mechanism for maintaining order. Cases like Dr Sinha,[50] who administered a lethal overdose of morphine to a patient with kidney failure, have proven the need for criminal sanctions for improper medical practice. Furthermore, in Garg,[51] which concerned a doctor who failed to recognise infection in his patient, the Court of Appeal confirmed that the legal framework of medical manslaughter must still reflect 'the fatal consequences of a criminal act.'[52]

This supports the argument that a revised criminal offence based on recklessness is necessary. This is because it proposes punishing doctors who deviate from the accepted standard of care while leaving some leeway for those who do not deserve to be held accountable to the same standard, which the current law falls short of doing. Therefore, this reform can be viewed as a positive step as it strives to promote justice and clarify a conducting legal framework. In addition, it will have a significantly more significant deterrent effect in discouraging poor and reckless practises by ensuring doctors consider their actions before carrying them out and weighing them against the risks to the patient. Punishing medical conduct that is genuinely inadequate and careless, rather than unfortunately fatal, will be far more effective in deterring irresponsible doctoring and encouraging careful practice and the avoidance of known mistakes.

Nonetheless, given the strain on healthcare resources in the UK, particularly during the COVID-19 pandemic, when egalitarianism gave way to utilitarianism and resources had to be rationed on a priority basis, the proposed reform is all desirable, practical, and effective. This is because it intends to protect innocent doctors who make mistakes under stress, enhance work systems, and promote patient safety without changing the healthcare infrastructure. Thus, the proposal advocated for in this paper can distinguish more successfully between those who are morally culpable and deserving of punishment and those who should be excused. Furthermore,

[49] Melinee Kazarian, Danielle Griffifths and Margaret Brazier, 'Criminal responsibility for medical malpractice in France' [2011] 27 *Professional Negligence* 185.
[50] BBC News, 'Doctor Jailed for Morphine Death' (2004) <http://news.bbc.co.uk/1/hi/england/merseyside/3601129.stm> accessed 20 January 2022.
[51] *Garg v R* [2012] EWCA Crim 2520.
[52] *R v Holton* [2010] EWCA Crim 934, (Lord Judge CJ).

it targets deliberate deviations from good practice instead of medical errors that have unintended repercussions.

Defining Recklessness

Concern has been expressed regarding the term 'recklessness' as it has confused legal practitioners. The Law Commission stated that despite this, 'there is no other word equally suitable to serve as a label for unreasonably taking the risk of which the defendant is aware.'[53] This may be partly because changing the offence by either adding or removing a word is not desirable as it is merely a label. Despite the efforts of many, the courts have not settled on a coherent doctrine and neglected the adoption of recklessness. However, this should not be a reason to turn a blind eye to recklessness, as Williams and Duff's definition strikes the correct balance for this, which should satisfy those wary of strict subjectivism.

Based on Williams and Duff's definition of recklessness, criminal liability can be attributed to a doctor having special knowledge that specific procedures pose certain dangers and failing to evaluate those risks without justification. This proposal based on 'practical indifference', which essentially means 'I could not care less', would best represent subjective recklessness to the CPS and the jury. Notedly, the Court of Appeal's endorsement of the trial judge's directive regarding the grossness of the defendant's conduct in *Misra* lays a strong foundation for such reform. It must be 'exceptionally bad', indicating an indifference to a serious risk to the patient's life. Thus, embracing indifference allows for mitigation in certain situations to reduce the severity of the law on those who are doing their best as it ensures the consideration of the defendant's state of mind. This supports the shift from negligence to recklessness as the proposed definition can distinguish between irresponsible doctors and doctors who make mistakes.

Deterring Dangerous Conduct

Gross negligence has been preferred over recklessness for deterrence purposes[54] as evidence shows a growing demand for healthcare practitioners to be held accountable for their mistakes.[55] Criminalising any conduct that ought to be dangerous, regardless of the outcome and any mitigating factors, have been favoured because it acts as a deterrent for the repetition of such actions, promoting safe practice and maintaining high standards of practice. Many academics and policymakers are concerned with 'forward-looking' grounds for criminalising and punishing behaviour; in essence, misconduct prevention is a significant priority. However, critics of negligence-based liability question whether it is feasible to prevent people from overlooking something they should have considered. Merry claims that medical errors that are not caused by subjective intent cannot be averted by rational thinking in most extreme situations. Hence, the prospect of criminal prosecution is ineffective.

[53] Law Commission, *Legislating the Criminal Code – INVOLUNTARY MANSLAUGHTER Item 11 of the Sixth Programme of Law Reform: Criminal Law* (Law Com No 237**,** 1965).
[54] Larry Alexander, Kimberly Kessler Ferzan and Stephen J. Morse, *Crime and Culpability: A Theory of Criminal Law* (Cambridge University Press, 2009).
[55] Melinee Kazarian, Danielle Griffifths and Margaret Brazier, 'Criminal responsibility for medical malpractice in France' [2011] 27 Professional Negligence, 185.

As Merry and Brookbanks have argued, the law is 'founded on a rejection of the nature of human error'[56] and punishes medical manslaughter harshly due to the acquired skill and knowledge.[57] Hart cast doubt on this notion as he questioned whether frequently telling a child to be careful is conceptually incoherent or consistently ineffective.[58] He states that medical professionals will learn to be more cautious if there is a threat of prosecution. Moreover, he adds that it is vital to remember that gross negligence, not simple negligence, is the legal threshold for liability as it is only the 'worst' cases considered.[59] Despite this, many instances deemed not to be 'gross,' such as that of Dr Tawana,[60] are deterrable; it is an example of where if the warning 'be more careful' had been heeded, lives would not have been lost.

Academics appear to agree that prosecutions may be counterproductive to excellent medical practice and obstruct future accountability. Excellent medical practice is built upon transparency and the opportunity to learn from mistakes. It has long been claimed that pointing the criminal finger of blame when harm is done is destructive to this.[61] More broadly, Ashworth observes that the deterrent power of prosecution is frequently overestimated, particularly in circumstances of medical error, where there is no wilful wrongdoing.[62] Currently, there is limited evidence that previous medical manslaughter prosecutions have improved patient safety or the systematic flaws that lead to deadly errors. For instance, despite the high-profile example of *Prentice* and *Sullman*, the fatal error of injecting vincristine into a patient's spine occurred again in 2001, resulting in death, totally the thirty-sixth instance of a deadly vincristine injection reported worldwide.[63] Thus, it should be acknowledged that medical errors, including fatal ones, are an inevitable feature of healthcare delivery. Therefore, the current law's deterrent effect as a way of maximising patient safety is hampered by the broad and arbitrary interpretation of the law and the fact that it applies to unintentional errors. Thus, this supports the argument for recklessness by refuting the much-prioritised necessity of gross negligence for retribution and deterrence. Assumptions should not be made that these are the only motivations for criminalisation as forward-thinking reasoning can aim to reduce crime using other methods. This essay agrees with Berwick that criminal sanctions should only be used in exceptional circumstances to deter irrational neglect and mistreatment.[64]

---

[56] Alan Merry and Alexander McCall Smith, '*Merry and McCall Smith's Errors, Medicine and the Law' (*Cambridge University Press, 2001), 104.
[57] Alan Forbes Merry, '*When are errors a crime? Lessons from New Zealand'* in Charles A Erin and Suzanne Ost (eds) *The Criminal Justice System and Health Care* (Oxford University Press 2007).
[58] H.L.A Hart, *Punishment and Responsibility: Essays in the Philosophy of Law* (2nd edn, Oxford, Oxford University Press, 2008).
[59] ibid.
[60] ibid., [26], 131.
[61] Oliver Quick, 'Patient safety and the problem and potential of law' [2012] 69 Cambridge Law Journal 186; Melinee Kazarian, Danielle Griffifths and Margaret Brazier, 'Criminal responsibility for medical malpractice in France' [2011] 27 *Professional Negligence* 185; Sarah E McDowell and Robin E Ferner, 'Medical manslaughter: Editorial' [2013] BMJ 347.
[62] Andrew Ashworth, 'Is the Criminal law a lost cause?' [2000] 116 Law Quarterly Review, 14.
[63] Margaret Brazier and Neil Allen '*Criminalising Medical Malpractice'*, in Charles A Erin and Suzanne Ost (eds), *The Criminal Justice System and Healthcare* (Oxford University Press, 2008), 23.
[64] National Advisory Group on the Safety of Patients in England, 'A Promise to learn – a commitment to act. Improving the safety of patients in England' [2013] <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/226703/Berwick_Report.pdf> accessed 09 January 2022.

Proving Subjective Recklessness

Quick argues that reckless manslaughter, based on a subjective test, would be a good substitute for gross negligence manslaughter. This is because it would set the level of liability at an appropriate level and offer greater certainty for prosecutors, judges and juries who currently struggle with the vague and imprecise notion of gross negligence. This would subsequently reduce the number of prosecutions against healthcare professionals.[65]

However, the Court of Appeal shows a preference for gross negligence due to concerns that proving subjective awareness of a risk would be difficult. This was partly due to the broad definition of recklessness which they claim causes difficulty for the ordinary lawyer and juror who may have felt that the word 'reckless' had stricter connotations, presumably accounting for capacity. The Court believed it had caused problems for cases of involuntary manslaughter that would not have happened had gross negligence been the test. Furthermore, a defendant's subjective awareness at the time of an incident is not easily proved. Healthcare professionals have special knowledge and should be aware of the risks of death in a particular situation does not mean that proof that they were aware could be proven without objective evidence, as they may merely ignore warnings. Without objective evidence, proof of gross negligence will be used, as it is in other cases, as evidence that the defendant must have realised what is being alleged. However, Quick cites a prosecutor who states, 'I can't see how we would bring a prosecution without an element of subjective recklessness… even if there's no direct evidence of subjective recklessness… but it may be so blindingly obvious that anyone must have realised….'[66] Yet, what is evident to a prosecutor working from a written file will often not be apparent to a jury following oral evidence and hearing the defendant. Instead of the jury determining what gross negligence means and whether it adequately depicts the defendant's behaviour, a subjective recklessness test will require the jury to decide whether the defendant understood what was apparent in hindsight.

Nonetheless, the few cases that have resulted in conviction are all situations of subjective recklessness. For example, despite warnings that removing a liver tumour was too risky, Dr Walker persisted. Consequently, he pleaded guilty.[67] Moreover, Dr Sinha, who provided a large morphine dosage to a patient with kidney failure to treat the pain of severe arthritis, had refused to view her medical chart, despite her husband's suggestion and was oblivious to the danger so consequently received a prison sentence.[68] Furthermore, Dr Ramnath, against the recommendation of three colleagues, administered a deadly amount of adrenaline, which was deemed a professional violation.[69] Therefore, the worry that prosecutors could not prove subjective awareness of risk is exaggerated. As evident, some prosecutors have

[65] ibid., [42], 202.
[66] ibid., 193.
[67] Owen Dyer, 'Surgeon accused of operating beyond his competence' [2001] 59 British Medical Association, 2.
[68] Liane Katz, 'GP Jailed for Giving Lethal Morphine Overdose' (*The Guardian*, 2004) <https://www.theguardian.com/uk/2004/apr/06/health.nhs> accessed 20 January 2022.
[69] David Wilkes, 'Doctor Killed Bunion Patient with Fatal Dose of Adrenaline, Court Hears' *Daily Mail* (2009) <https://www.dailymail.co.uk/news/article-1114606/Doctor-killed-bunion-patient-fatal-dose-adrenaline-court-hears.html> accessed 20 January 2022.

been searching for evidence of subjective fault, perhaps in agreement with Williams that recklessness is 'socially superior' to gross negligence.[70]

Incompetence in Eliminating Foreseen Risks

The Court of Appeal asserted that gross negligence is sufficient as recklessness does not cover a situation where a doctor has foreseen a risk and tries to avoid it but does so in an incompetent manner.[71] If the doctor has eliminated the risk before acting, they would still be caught by the test for gross negligence, leaving the jury with a vague term and a complex legal question to answer. Yet, they could equally have been deemed reckless by closing their mind to the fact that the risk remained.

Under the recklessness approach, it is possible that in such circumstances, the defendant could be deemed reckless in choosing to eliminate or avoid the risk in the way they did.[72] Alternatively, they could be simply negligent because they did their incompetent best and honestly thought he had done enough to prevent harm.[73] In the latter case, criminal liability would not be justified. However, Quick states that the concept of negligence also entails significant problems, and such cases as those described above should 'fall below the line' and not be held criminally liable.[74] He adds that the two most prominent arguments against this do not appear strong. Firstly, no one has persuasively evidenced how more prosecutions in healthcare would be beneficial, for instance, in improving patient safety. As argued earlier, deterrence in prosecuting medical practitioners has been proved ineffective, so there is no rationale behind holding every case of harm criminal. Secondly, it may be questioned whether the shift from gross negligence to recklessness would make any difference in practice, given their similar nature.[75] Since demonstrating recklessness is more difficult for the CPS, it would undoubtedly result in fewer instances going to trial. Thus, recklessness is a desirable shift from negligence. There are clear guidelines in how to approach 'borderline' cases where a criminal sanction would sometimes be applied unjustly under the law of negligence. Under recklessness, considering the defendant's state of mind, even if they cannot eliminate the risk once aware, it is unjust for them to be held criminally liable for their inability.

Conclusion

Medical errors are inevitable in healthcare, rendering the current legal framework of medical manslaughter unsatisfactory and ineffective due to its ignorance of mistakes. Many concerns around medical manslaughter stem from the lack of legal clarity, as gross negligence is too vague due to the common law's failure to address the courts' circular statements about what constitutes 'gross.' While academics seek clarity from higher courts on the elusive nature of this term, Lord Mackay's dismal circular answer in *Adomako* provides the only insight. Despite the Court of Appeal's affirmation of the *Adomako* test's certainty in *Misra*, the law's application in courts and at the prosecution stage has remained unclear and inconsistent.

[70] Glanville Williams, *Textbook of criminal law* (2nd edn, London, 1983), 100.
[71] ibid., [15].
[72] ibid., [37].
[73] ibid., [37].
[74] ibid., [42].
[75] ibid., [42].

There is widespread dissatisfaction amongst prosecutors and academics with regards to the inability of the law to differentiate between incompetent doctors and those who make momentary mistakes. Because the law disregards defendants' moral culpability, it is overly harsh on those defendants who make unintentional errors when under pressure and trying to do their best for their patients. This raises doubt that there is insufficient moral blame for criminal liability. While the purpose of this paper was not to argue that medical professionals should be given special treatment under the law, it did lay out a framework that allows for flexibility to accommodate the demanding circumstances under which doctors work. As argued by Quick, there is no feasibility in continuing the struggle of interpreting gross negligence when there is no consensus as to what it means and whether it should be a crime.

Recklessness has been labelled as 'socially superior' to gross negligence as a means of determining liability,[76] which is defensible given the improvements that the proposed test will potentially bring to both healthcare and justice. Within the test, the centrality of 'indifference' ensures the consideration of the defendant's state of mind, promoting diligent practises when making healthcare decisions. Furthermore, it allows for mitigation in certain situations to reduce the severity of the law on those who are doing their best despite their circumstances. Arguably, enforcing the proposed reform will improve patient safety by deterring poor healthcare while leaving leeway for human error. It is critical that any healthcare reform acknowledges the inevitability of error and uses it as a mechanism for improvement rather than punishment. Thus, criminal law should be used only as a last resort and rarely. As such, the appropriate place to set liability should be recklessness. While this concept is not without difficulties, reformist energies should be channelled towards its revision. Such a reform of this nature is necessary for promoting diligent healthcare and aiding the pursuit of justice by clarifying the substantive law with a stronger focus on culpability.

[76] ibid., [54].

Meet BCLP:
Our Firm, Our Recruitment

BCLP is an international law firm with 1,400 lawyers across 30 offices worldwide. Working as one team, we bring to our clients a range of integrated capabilities, including some of the world's most active M&A, real estate, financial services, litigation and investigations practices. Join us and you'll work on the most complex legal matters, finding the kind of cutting-edge solutions that have seen us win a host of awards.

MEET YOUR FUTURE

bclplaw.com

#MEETBCLP

Defining Trans Parenthood - *McConnell v Registrar General for England and Wales*

Tina Ye

Introduction

The respect for private and family life remains one of the most important human rights protected by the European Convention of Human Rights (ECHR). As the English Court of Appeal in *R (Elan-Cane) v Home Secretary*[1] summed up nicely, 'there can be little more central to a citizen's private life than gender, and gender is one of the most important aspects of private life.' However, despite this endorsement by the court, transgender applicants still face precarious chances when bringing these claims to court.

In *R (on the application of McConnell) v Registrar General for England and Wales,*[2] a transgendered man wished to be registered as a 'father' or 'parent' on his child's birth certificate but was denied due to exceptions stipulated in Section 12 of the Gender Recognition Act 2004 (GRA).

Facts of the Case

The claimant, a transgendered man, underwent gender affirming surgeries and obtained his Gender Recognition Certificate (GRC) in 2017. The certificate, described under Section 9(1) of the GRA, lends legal legitimacy to the person's acquired gender 'for all purposes.' In the same year, he underwent intrauterine insemination treatment in hopes to carry a pregnancy. In January 2018, his son was born. Upon registering for his son's birth certificate, he was informed that he would have to be registered as the child's 'mother' pursuant to Section 33 of the Human Fertilisation and Embryology Act 2008 (HFEA), even though his legal gender is male on all legal documents.

The claimant sought judicial review against the registrar's decision on two grounds. First, he claimed that he should be entitled to register as 'father' because the GRC recognised his acquired gender for all legal purposes. Second, if he was not entitled to do so, then a declaration of incompatibility should be issued under Article 8 and Article 14 of the ECHR for discrimination against his rights to respect for family and private life.[3]

The Ruling

The Court of First Instance rejected his claims. In a judgement delivered by Sir Andrew MacFarlane, it was observed that Section 9 GRA stipulated exceptions to GRC's legal recognition under certain circumstances and that such exceptions applied to parenthood. The term 'mother' is to indicate one's 'biological role in giving birth', irrespective of one's acquired

[1] [2020] 3 WLR 386, [46].
[2] [2020] EWCA Civ 599.
[3] ibid., [10].

gender.[4] Sir MacFarlane did acknowledge that there was an interference with the claimant's Article 8 rights,[5] but he reasoned that the interferences with the claimant's Article 8 rights were necessary and proportionate. It was accepted that the need for a coherent and certain birth registration scheme was necessary,[6] and it must be in the best interest of the child.[7]

Upon appeal, the Court of Appeal unanimously dismissed the appeal. The judgement first examined the interpretation of Section 12 GRA (entitled 'Parenthood') and its effect on Section 9 GRA. Section 12 reads '[t]he fact that a person's gender has become the acquired gender under this Act does not affect the status of the person as the father or mother of a child.' The appellant asserted that this only had retrospective effect, meaning it bound him to his birth gender if he had his son prior to him obtaining the GRC. The court sided with the respondent and recognised that it was both retrospective and prospective.[8] This was done by interpretation in its ordinary meaning, compared with wording in other sections of the GRA stipulating exceptions to Section 9 and pinpointing express language used in sections which were intended to only have retrospective effect. The court concluded that Section 12 is one of the exceptions set out in Section 9.

Established thus, the court then assessed if these provisions are compatible with the ECHR articles. Echoing the reasons put forth by the First Instance judge, the court held that although there were interferences with the appellant's ECHR rights, the interference was 'in accordance with the law' and the interference had a 'legitimate aim.'[9] Additonally, there were no less intrusive means available to alter the scheme and a multitude of legislations would be affected should the word 'mother' be changed in its interpretation. Moreover, there was no Strasbourg jurisprudence on similar cases which left the court to confer a wider margin of judgement to Parliament.[10] For all of the reasons above, the court concluded that there was no incompatibility between the provisions and ECHR.

Commentary

With the influx of transgender identity cases brought before the court, the judiciary must weigh their decisions in this rapidly developing area of law carefully. The reluctance to take a broad stance can be explained by a fear of overstepping and undermining Parliamentary sovereignty. As Sarah Williams, Head of Family Law at Payne Hicks Beach, writes: '[the court's] role is to react to new concepts rather than create them.'[11] The Court of Appeal offered a 'wide margin of judgement' to Parliament, acknowledging its comparative democratic legitimacy to intervene on matters of 'controversial social policy.'[12] While this judicial restraint can be lauded in its pursuit to maintain the balance of power and prevent acts of 'legislating from the bench', it conveniently allowed the court to avoid addressing complex underlying issues of

[4] ibid., [279].
[5] ibid., [250].
[6] ibid., [263].
[7] ibid., [262], [263].
[8] ibid., [29].
[9] ibid., [57]-[58].
[10] ibid., [79]
[11] Sarah Williams, 'Trans-Parenthood and a Minor's Ability to Consent to Gender-Changing Medical Treatment' [2022] ICLG Family Law 5, 4.
[12] *McConnell* (n 2), [82].

transgender identity and parenthood. Alan Brown, writing in the context of surrogacy, argued that 'the promotion and protection of the traditional nuclear family was influential in shaping [the] initial legislative approach.'[13] This can, and should, be aptly applied to the current case as the intent to maintain traditional family values pervaded through the legislative approach on legal determinations of family law. By focusing intensely on statutory interpretation, the court took the narrative away from transgender familyhood and undermined the much-needed challenge to the status quo in English family law.

Perhaps, the courts' view on gender is the clearest indicator of the its oversight whilst adjudicating on transgender parenting. The High Court and Court of Appeal both stated that 'mother' is not tied to a single gender, but rather describes an individual who 'undergoes the physical and biological process of carrying a pregnancy and giving birth.'[14] The explanation was expanded further to say that the law recognises mothers who are biologically male and fathers who are biologically female, thereby actually *embracing* a progressive attitude on transgender parenthood. This explanation is unconvincing and represents a blindness to the lived reality of transgender parents. To define 'mother' and 'father' as genderless terms embodies a stance incontrovertibly removed from the collective reality of everyday life. Mothers and fathers have always been traditionally assigned to respective genders and its social connotations are clear no matter what the law tries to relay. This is especially disorientating to transgender parents; to require a transgendered man to identify as a 'mother' on a formal public document could only exacerbate their gender dysphoria and cause lasting psychological ramifications.

This was not made better by the court's response to the appellant's contention that the legislation should be construed in such a manner as to reflect current social modes. Citing Lord Bingham in *R (Quintavalle) v Secretary of State for Health,*[15] the court stated that 'if and in so far as the argument is that the word "mother" should be construed as "father", that would offend against the principle [...] that the word "dog" cannot be construed to mean "cat".'[16] It may be just a simple repetition of the language used in a leading authority but it, in making the comparison of gendered terms like mother and father to dog and cats, speaks loudly at the malicious ignorance about gender that pervades conservative society. As Alan Brown comments, 'the use of this language could be seen as either providing evidence of the Court of Appeal's privileged carelessness or as being deliberately provocative.'[17]

Another interesting point to note is the multiple references made to the child's interest. In the High Court, it was acknowledged that although it would cause the parent and the child embarrassment and distress if the child's full birth certificate (which contains the mis-gendered identity of the parent) was presented, it was more important to strike a balance between the parent's right to privacy and the child's right to know about their parent's biological identity.

[13] Alan Brown, 'Two Means Two, but Must Does Not Mean Must: An Analysis of Recent Decisions on the Conditions for Parental Orders in Surrogacy' [2018] Child and Family Law Quarterly 30, 23.
[14] *McConnell* (n 2), [35].
[15] [2003] UKHL 13, [9].
[16] *McConnell* (n 2), [35].
[17] Alan Brown, 'Trans Parenthood and the Meaning of "Mother", "Father" and "Parent"' [2021] Medical Law Review 29, 157.

The court cited *Godelli v Italy* (unreported),[18] a case where the claimant wished to find her biological mother's identity but was unable to do so because her mother had assumed anonymity on her birth certificate, and this caused the claimant psychological suffering. The biggest difference between *Godelli* and transgender parenthood cases like *McConnell* is that the latter does not wish for anonymity — he is not hiding his 'true' identity nor is he lying to his child. He simply wished to have his lived reality reflected in his private and public life.

It must be addressed that by making a comparison between two disparate cases highlights the harmful assumptions underlying such comparisons and it is these very assumptions which continue to fuel the stigmatisation of transgender people today. The first assumption is that transgendered parents would keep their identity hidden from their child, out of shame or out of a personal selfishness. This assumption plays at the core of transphobia: nonconformity to cis-heteronormality must be concealed — it is a deviancy and outside the remit of civilized society. It plays at the assumption that transgender parents would not divulge their identities, willingly, unless ordered by law. Therefore, following a sense of judicial paternalism, the law must ensure the execution of such an act.

The second assumption is that the child would suffer psychological distress if they were denied this right to know. There is a certain urgency evoked by the insistence on the child's wellbeing, and most often, rightfully so. Children are a vulnerable section of society and should be protected. However, in the context of transgender parenthood, when there is declaration of a balancing exercise,[19] it can favour the scale towards one end over the over. The High Court judgement expressed this sentiment - fearing that, without such protection for the child, the child would be 'marked out' from all other children. This effectively villainises the transgendered parent (i.e., as if to say 'why would you do this to your child') and victimises, unfairly, the child (to say 'you should not have to suffer like this' when the suffering is purely imagined and based on prejudice). It ignores the realities that transgendered parents feel the same love for their children and that a parent-child relationship *not* founded on the nuclear family model could work just as well. To bring up a powerful quote, Professor Golombo writes - quite elegantly and with impact - in her book, '[o]ur research so far has shown that changing identity does not preclude parents from being protective of, and loving towards their children, and neither does it cause children to develop psychological problems. Despite the hurdles they face, children seem to adapt to their parent's transition.'[20]

By claiming to act in the child's best interest, the court engages in judicial paternalism. The court felt justified in protecting the child's interest against that of their parent, overlooking the fact that transgendered parents should have a choice in how they wish to reveal their biological identity to their child — a choice that is contingent on a respect for private family life. Perhaps this insistence to watch out for the child's wellbeing is once again a condemnation of the unconventional family structure. As Claire Fenton-Glynn writes, 'the law prioritises the conventional family, and marginalises those who fall outside [of it].'[21]

[18] Godelli v Italy CE:ECHR:2012:0925JUD003378309
[19] [2019] EWHC 2384, [250-258]; *also see* [2020] EWCA Civ 559 [85-88].
[20] Professor Susan Golombok, *We are Family: The Modern Transformation of Parents and Children* (Scribe 2020).
[21] Claire Fenton-Glynn, 'Deconstructing parenthood: what makes a "mother"?' [2020] Cambridge Law Journal 79, 34.

One can argue that the courts approach to transgender issues will always be one step behind, if not altogether prejudiced. Judicial attitude on transgender familyhood is unfortunately stuck in the past. Ten years ago, McCandless and Sheldon wrote: 'the questions posed by transgender parenthood serve to illuminate many of the tensions inherent in continuing our legal determinations of parenthood to a family model that is unmoored from its traditional underpinnings.'[22] However, it seems that *McConnell* only illuminated the stagnancy and opaque reflection of transgender parenthood as it continues to exist in the law to this day. Until the court is willing to step away from technical narratives and step into (if not embrace) a non-traditional, non-conformist, non-heteronormative approach to family creation, the lived realities of many transgender families will continue to be obscured.

[22] Julie McCandless and Sally Sheldon, "The Human Fertilisation and Embryology Act [2008] and the Tenacity of the Sexual Family Form" [2010] Modern Law Review 73, 175.

The Prejudicial Effect of the Past

Sophia Evans

Abstract

The prejudicial effect of admitting evidence of a defendant's past behaviour requires the law to be cautious, particularly regarding evidence of bad character and evidence of a complainant's sexual history. A jury in a criminal trial may otherwise too readily conclude that a person has acted in a manner consistent with his or her past behaviour, potentially resulting in a miscarriage of justice.

This article aims to critically assess why the law must be cautious about allowing the admission of evidence of defendants' bad character and evidence of complainants' sexual history. In order to provide the necessary context, this essay will begin by reviewing the development of the law and how the law is applied today. This will be followed by an evaluation of arguments concerning the extent to which such evidence is affected by the alleged prejudice of the jury, concluding that the law must never neglect to consider the possibility of such prejudice as it may bear considerable effect on the issue of guilt and outcome of a trial.

The Admissibility of Bad Character Evidence

Evidence of bad character is subject to a number of assumptions regarding its relevance. It is assumed that such evidence is relevant to propensity and credibility. Plainly stated, 'evidence of a person's bad character is considered relevant to his propensity to act in a manner consistent with that bad character, as well as relevant to his credibility.'[1] It follows that in a criminal case, a jury may consider evidence of bad character relevant to the issue of guilt for the offence with which he is currently charged. This assumption leads to the risk of two types of prejudice, known as 'reasoning prejudice' and 'moral prejudice.' The first arises where a jury may 'overestimate the probative value of the evidence in question', whereas the second arises where a jury may 'return a guilty verdict because (the defendant) is thought to deserve punishment for previous misconduct.'[2] Research on magistrates shows that 'any previous conviction, recent or old, affected … the defendants likely guilt and verdicts unfavourably.'[3,4] This is supported by the findings of the LSE Jury Project, thereby furthering this point.[5]

The admissibility of bad character evidence was governed by old common law principles until replaced by the Criminal Justice Act 2003.[6] These provisions generated many critiques and were, notably, not well received by the Court of Appeal.[7] The provisions

[1] Andrew Choo, *Evidence* (5th edn. Oxford University Press 2018).
[2] ibid..
[3] Sally Lloyd-Bostock, 'The Effects on Lay Magistrates of Hearing that the Defendant is of "Good Character" Being Left to Speculate. Or Hearing that he has a Previous Conviction' [2006] Criminal Law Review 189, 211.
[4] It should be noted that recent convictions which are recent and dissimilar to the current charge are excluded from this finding.
[5] LSE Jury Project, 'Juries and the Rules of Evidence' [1973] Criminal Law Review, 208.
[6] Criminal Justice Act 2003, s.99(1).
[7] *R v Bradley* [2005] EWCA Crim 20, 1 Cr App R 24, [39].

implemented by this Act stand alongside other statutory provisions such as s.27(3) of the Theft Act 1968 for which there have been 'a number of calls for repeal … including a recommendation to that effect from the Law Commission.'[8] The CJA 2003 defines bad character evidence as 'evidence of, or of a disposition towards, misconduct', misconduct being defined as 'the commission of an offence or other reprehensible behaviour.'[9] The admissibility of bad character evidence of a non-defendant is outlined in s.100 CJA 2003. More importantly, the admissibility of bad character evidence of a defendant is outlined in s.101(1) CJA 2003, requiring one of 'the seven gateways' to be satisfied before evidence can be used. The fourth gateway (d) is especially noteworthy as it 'replaces the 'similar fact' rule at common law with a principle whereby evidence of a defendant's bad character is prima facie admissible on the issue of guilt so long as it is relevant.'[10] The considerably large volume of case law generated in the wake of these provisions being implemented in 2005 provides two general observations. The Court of Appeal case of *R v Hanson* says, 'the purpose in legislation … was to assist in the evidence-based conviction of the guilty, without putting those who are not guilty at risk of conviction by prejudice.'[11] Secondly, also in the Court of Appeal, the case of *R v Campbell* notes 'once evidence has been admitted through a gateway it is open to the jury to attach significance to it in any respect which is relevant.'[12]

These interpretations demonstrate the 'common sense' approach adopted by the courts and the considerable amount of confidence placed in the discretion of trial judges. It is, therefore, left to them to strike a balance between freer admissibility of evidence and caution to be exercised when dealing with evidence of bad character. This approach is highly optimistic, especially in light of the 'unacceptable risk of prejudice to accused persons and potential for miscarriages of justice' born from the provisions in the CJA 2003.[13] It is for this reason that we can conclude that courts should exercise extreme caution when allowing evidence of bad character to be admitted for fear a jury may, too readily, conclude a person has acted in a manner consistent with his or her past behaviour.

The Admissibility of Sexual History Evidence

The same risk of prejudice examined above exists in the admissibility of evidence of a complainant's sexual history. The dangers of admitting such evidence are, therefore, not dissimilar with regards to relevance. Historically, evidence of sexual history was readily admitted before the court in order to infer consent and challenge credibility. Early case law demonstrates, what is now known as, the 'twin myths'; that 'sexually active women are less credible as witnesses and more likely to consent.'[14] The discreditation of such inferences sparked the first wave of reform to restrict the admissibility of sexual history evidence; such as

[8] ibid at 1; Law Commission (Law Com No 273) *Evidence of Bad Character in Criminal Proceedings* [2001] [11.55].
[9] Criminal Justice Act 2003, s.98 and s.112(1); It should be noted that while 'reprehensible behaviour' has no statutory definitions, the court of appeal case *R v Renda* [2005] provides that 'as a matter of ordinary language, the word 'reprehensible' carries with it some element of culpability or blameworthiness.'
[10] ibid., at 1.
[11] *R v Hanson* [2005] EWCA Crim 824, [4].
[12] *R v Campbell* [2007] EWCA Crim 1472 [25], [28].
[13] ibid.
[14] *R v Seaboyer* [1991] 2 SCR 577 [634].

the recommendations of the 1975 Heilbron Report.[15] The Sexual Offences (Amendment) Act 1976 placed 'faith in judicial discretion to limit the admission of such evidence' however in effect 'did little to stem the flow of sexual history evidence being admitted: once again leading to demands for legislative reform' according to Lord Steyn in the case of *R v A.*[16,17] It is for this reason that the Youth Justice and Criminal Evidence Act 1999 was adopted, implementing an exhaustive list of exceptions to the general rule of exclusion thereby abolishing the ineffective discretionary approach of the court. 'Despite clear legislative intent to curtail the use of sexual history evidence', further proposals for reform made their way to Parliament in the wake of the controversial judgement of *R v Ched Evans,* a case concerning rape and third-party evidence. The Court of Appeal held that 'sexual history evidence relating to persons other than the accused was admissible and potentially relevant.'[18] It should be noted that, the defendant obtained an acquittal in the retrial where this evidence was introduced.

The current law is set out in s.41 of the YJCE Act 1999, which provides that no question or evidence about any 'sexual behaviour' of the complainant may be adduced at trial. The term 'sexual behaviour' is defined simply as 'any sexual behaviour or other sexual experience' and applies to evidence relating to the sexual activity of both the accused and third parties.[19] It should be noted that s.41 does not provide for any overriding judicial discretion to admit evidence, suggesting a small victory for reform in light of the problems arising from the Sexual Offences (Amendment) Act 1976. The restrictions in s.41 only apply to the defence, leave for which may only be granted under the prescribed four exceptions. These include evidence relating to a relevant issue in the case and is not an issue of consent, evidence relating to sexual behaviour at or about the same time of the sexual activity in question, evidence which is 'so similar that the similarity cannot be reasonably explained by coincidence' and evidence necessary to rebut prosecution claims.[20] The 2001 House of Lords case of *R v A* presents a significant legal challenge to the restrictions set out under s.41 on the basis that these provisions infringed a defendant's right to a fair trial under article 6 of the Human Rights Act 1998. Ultimately, the courts took the view in the case *R v Hamadi* that 'the wider importance of *R v A* lies in the recognition that protecting complainants from indignity and humiliating questions to which s.41 is directed, must ultimately give way to the right of fair trial.'[21] The judgement given in this case is one of many highlighting the ambiguities in this area of law. The most recent of these being the *Evans* case, after which the government announced that a review of s.41 is currently underway.[22]

[15] Heilbron Committee, *Report of the Advisory Group on the Law of Rape*, Cmnd 6352 (HM Government, 1975).
[16] Clare McGlynn, 'Rape Trials and Sexual History Evidence: Reforming the Law on Third-Party Evidence' [2017] 81(5) Journal of Criminal Law, 367.
[17] *R v A* [2001] UKHL 25, [28].
[18] *R v Ched Evans* [2016] EWCA Crim 452.
[19] s.41(1)(c); *R v Mukadi* [2003] EWCA Crim 3765, [14]: '[definition of sexual behavior] is really a matter of impression and common sense.'
[20] S.41(3)(a)(b)(c) and (5).
[21] *R v Hamadi* [2007] EWCA Crim 3048, [18]; ibid., [18].
[22] Nick Lester, 'Government to Review the Law of Protection of Rape Complainants in the Light of the Ched Evans Case' *The Independent* (17 November 2017); ibid., [20].

While one may consider this as a formal acknowledgment of the need for reform, it can be argued that 'much of the damage to the public perception of how the criminal justice system deals with allegations of serious sexual offences in general... arose not from the decision of the Court of Appeal in *Evans*, but from the reporting of it.'[23] Furthermore, the Attorney-General's Office and the Ministry of Justice in its review of applications to admit evidence of a complainants sexual behaviour considers the cases of *Mukadi* and *Evans* to be 'outliers,' suggesting that s.41 of the YJ&CEA 1999 is operating as intended.'[24] Inapplicable to the normative operation of s.41, it would therefore follow that a call for the wholesome reform of provisions may be considered unnecessary. According the recent research by Laura Hoyano at Oxford University, 'it is the complexity of s.41 that is problematic rather than the standard at which sexual behaviour evidence is considered sufficiently relevant to warrant its admission.'[25] Nevertheless, the prejudicial effect of sexual history evidence in the context of today is seen from evidence introduced as attacks on moral credibility 'showing the complainant to be so morally inferior as either not to deserve the court's sympathy or not to provide a suitable foundation for punishing the accused.'[26] It is for this reason, that 'sexual history evidence risks introducing irrelevant or prejudicial material which may distort rather than promote the truth-finding role of the trial and rectitude in its decision making.'[27] We can therefore conclude that courts should exercise extreme caution when allowing evidence of sexual history to admitted for fear a jury may too readily conclude a person has acted in a manner consistent with his or her past behaviour.

Conclusion

The principles of autonomy and free choice should not be infringed for fear of prejudice on the part of the jury, nor restrict a trial judge from delivering effective direction. This essay does not aim to discredit the grounds for the admissibility of all evidence pertaining to bad character or evidence of a complainant's sexual history. Instead, it aims to emphasise the importance of courts exercising extreme caution when doing so. A jury may otherwise misinterpret judicial direction, for at the end of the day, in a criminal trial, we are judged by a jury of our peers who are to be considered as laymen. Relying on judicial discretion, however, has also proved to be an undesirable solution. This is evidenced by the implementation of the Sexual Offences (Amendment) Act 1976 and addressed in issues arising from the great faith placed in trial judges regarding the admissibility of bad character evidence. Where exactly the threshold for the admissibility of bad character evidence and evidence of sexual history should

[23] Brian Brewis and Adam Jackson, 'Sexual Behavior Evidence and Evidence of Bad Character in Sexual Offence Proceedings: Proposing a Combined Admissibility Framework' [2020] 84(1) The *Journal of Criminal Law* 49-73.
[24] ibid.
[25] Laura Hoyano, The Operation of the Youth Justice & Criminal Evidence Act 1999 Section 41 in the Courts of England & Wales: Views from the Barristers' Row, [2019] Oxford Legal Studies Research Paper, 17.
[26] Aileen McColgan, 'Common Law and the Relevance of Sexual History Evidence' [1996] 16 OJLS 275. Note also that while it will rarely be declared that sexual history challenges truthfulness, sexual history evidence is admitted in order to demonstrate 'motive to lie' (see further below).
[27] ibid., at 18.

be will undoubtedly continue to be a matter of debate.[28] The combined evaluation of the prejudicial effect resulting from the admissibility of such evidence leads to the following conclusion: by creating a certain degree of consistency across the admissibility requirements in terms of relevance *in general* would likely result in a far less complex process. A principled regime throughout the law of evidence is, as the final point of this essay, long overdue.

[28] Matt Thomason 'Previous Sexual History Evidence: A Gloss on Relevance and Relationship Evidence' [2018] 22 The International Journal of Evidence and Proof 4, 342.

The Canadian Answer to the English Dilemma: Contribution-based Approaches over a Common Intention Constructive Trust in Family Homes

Alexander Clevewood

Abstract

Determining each other's proprietary interests in the shared home after relationship breakdown remains hotly debated. Common intention constructive trusts are currently used in the English legal framework. However, there are practical difficulties associated with finding a common intention. Where one could not be found, it may be imputed to the litigating parties. In *Stack v Dowden*, Baroness Hale stated that this was possible to achieve fairness. *Jones v Kernott* clarified that imputation could only operate in the quantification of interests. Also, imputation is limited to the (1) presumption of resulting trusts, or (2) absence of evidence to infer common intention. This clarification limited its use in legal practice, but imputation remains controversial since it practises the court's standard of morality.

However, this can be resolved by focusing on how the defendant is benefited, rather than summoning a common intention. This is particularly important when the defendant is the sole legal owner. After *Pettkus v Becker*, Canada shifted from an intention-based resulting trust to a remedial constructive trust approach, based on unjust enrichment. This shift absolves the need to find or impute a common intention. Any of the claimant's contribution received by the defendant in relation to the home, domestic or financial, already constitutes enrichment. Gender discrimination is avoided since the claimant's contributions are not judged against a (gendered) threshold for detrimental reliance. There has been academic discussion regarding whether this legal avenue itself leads to additional problems concerning the 'enrichment' itself. In the case of services rendered by the claimant as contribution to the family home, the defendant can 'subjectively devalue' such services. They are, therefore, not enrichment and do not attract compensation. However, subjective devaluation would normally be unavailable if it were unconscionable to do so, provided that the claimant rendered the service in expectation of a reward, and the defendant knew this. This would necessitate a convoluted inquiry of subjective opinions, hence complicating the law. The point of adopting an unjust enrichment approach is frustrating. This view is appreciated, and slight doctrinal modifications can be adopted to fill this legal lacuna.

This approach may also attract other sources of criticism. Remedial constructive trusts are rejected in *FHR European Ventures LLP v Cedar Capital Partners LLC* by the Supreme Court. Therefore, importing them into English law may lead to judicial uncertainty. However, any remedial constructive trust imposed would be based on an established English legal doctrine, instead of the adjudicator's ideals – the main concern underlying judicial rejection of this form of constructive trusts. Routes to reclaim assets discussed in this article include proprietary remedies from unjust enrichment, and conventional trusts law (resulting trusts and *Quistclose* trusts). Shifting one's attention away from the common intention is expected to promote more fairness in asset division.

Introduction

Common intention constructive trusts are the current English approach in determining shares in the family home after the separation of common law partners. However, this approach is marred with difficulties, and often clashes with notions of normative justice, particularly (1) the nebulous nature of common intention; (2) the practical application of imputation; and (3) compromised gender equality due to the operation of the detrimental reliance limb. A better solution should be sought. This article focuses on the Canadian approach, which entails the imposition of a remedial constructive trust based on unjust enrichment, established in the leading case of *Pettkus v Becker*. This article proceeds to discuss the feasibility of introducing this novel application of the doctrine of unjust enrichment to England and Wales, with a particular focus on the (1) 'enrichment' limb; and (2) the granting of proprietary remedies. This article also discusses the feasibility of a Quistclose trust / resulting trust analysis.

Current State of Family Homes - Common Intention Constructive Trusts

In *Stack v Dowden*[1], the House of Lords proclaimed that the use of resulting trusts in the familial context was no longer sound. In equity, resulting trusts were traditionally used since there was an ongoing presumption (imputed) that the donor, upon giving property to the donee, did not intend to relinquish all interests. This presumption was subject to exceptions, such as relationships of trust and confidence.[2]

Common intention constructive trusts (CICT) are an innovative approach adopted by courts subsequent to *Stack v Dowden*, departing from the resulting trust analysis. This is not limited to domestic scenarios, as erroneously envisaged by Lord Neuberger.[3] Moreover, after *Marr v Collie*[4], it is clarified that common intentions in relation to asset division, of litigating parties should be the priority given effect by the courts, regardless of scenario. The applicability of resulting trusts is accordingly vastly reduced in the family home context.

There are three stages in determining a claimant's proprietary interest under a CICT. If the claimant is not a legal owner: (1) Starting point (equity follows the law, meaning the claimant has no interest, both in equity and at law); (2) Acquisition stage (demonstration of express agreement, and detrimental reliance), and (3) Quantification stage. However, there are inherent issues in this approach. In the following, three problems are elucidated: (1) the practical difficulty of finding a common intention; (2) imputation of a common intention; and (3) gender discrimination flowing from the application of the detrimental reliance limb.

Practical Difficulty of Finding Common Intention

In uncomplicated cases of relationship breakdown, where trust and confidence are maintained throughout, finding a common intention is normally straightforward. In this ideal scenario, both parties have frankly expressed their opinions on more profound topics of relationship-related financial pooling, such as (1) shares in the family home; and (2) any financial contributions

[1] *Stack v Dowden* [2007] UKHL 17.
[2] I am referring to the use of the presumption of advancement, which is abolished by s.199 of Equality Act 2010.
[3] *Laskar v Laskar* [2008] EWCA Civ 347; 'In other words this was a purchase which, at least primarily, was not in 'the domestic consumer context' but in a commercial context. To my mind it would not be right to apply the reasoning in Stack v Dowden to such a case as this'; [17].
[4] *Marr v Collie* [2017] UKPC 17.

towards the acquisition/maintenance of the property. However, reality is unfortunately, very rarely straightforward. Deceit is rife in relationships. In *Grant v Edwards*,[5] the legal title of the contested property was vested in the defendant. The defendant was found to have told the claimant that doing otherwise would 'cause some prejudice in the matrimonial proceedings between [her] and her husband.' In *Eves v Eves*,[6] the property was again in the sole name of the defendant. The defendant alleged that 'as she was under 21, it could not be in joint names and had to be in his name alone.' He later clarified in the proceedings that it was an excuse. In both cases, the courts *expressly found* a common intention to share the property, despite the 'smokescreen lies' described above. Lord Denning, in *Eves v Eves*, stated that the defendant should be 'judged by what he told her -- by what he led her to believe -- and not by his own intent which he kept to himself.'[7] This alludes to an objective approach. On one reading, this is demonstrative of judicial clarity. Both parties *did* have a common intention to share the property together, but this intention was not manifested on paper due to operational difficulties and practical concerns, e.g., age limitations. The deceit was a subjective element, a figment of the defendant's unpronounced scheme. However, a second reading of the judgement produces uncomfortable results. As Gardner rightly pointed out, proffering excuses 'does not mean the men [the defendants] were thereby acknowledging an agreement whereby the woman should have a share.'[8] The excuse proves that there is no common intention to start with. It merely masks the actual reason behind the defendant's decision, which is immaterial owing to its subjective nature, as rightly suggested by Lord Denning.

In the author's opinion, logically, the second reading is more convincing. Returning to Lord Denning's pronouncement in *Eves v Eves*,[9] the defendant clearly denied the claimant a share in the property. The strength of the negative response is *softened* – but by no means *eliminated*, let alone supplanted by a positive one – by the fabricated excuse, which is understandably aimed at preserving the romantic association. An *actual* scenario of practical difficulties obstructing the realisation of a common intention (hence permitting the court to find a common intention accordingly) is not one involving the telling of fabricated excuses by one of the litigating parties, but one affected by the involvement of governmental authorities, or indeed any other source of third-party involvement which is not reasonably foreseeable by either party. If the involvement were reasonably foreseeable, this would take us back to the position above: whether the defendant is seeking to 'soften' the negative response and has never intended to benefit the claimant. Returning to *Eves v Eves*, evidence could be adduced to show that (1) the parties have attempted to implement their common intention, e.g., making official inquiries as to whether the property can be jointly held at law; and (2) regulatory or bureaucratic hurdles which prevent the common intention from being implemented, e.g., statutory restrictions on the legal age to hold property legally. Therefore, this creates doctrinal discomfort in decisions exemplified by *Grant v Edwards*, and *Eves v Eves*. More significantly, it highlights the practical infeasibility and considerable heterogeneity in finding a common

[5] *Grant v Edwards* [1986] Ch 638.
[6] *Eves v Eves* [1975] 3 All ER 768.
[7] ibid., 772b.
[8] Simon Gardner 'Rethinking family property' [1993] 109 LQR 1, 265.
[9] ibid., n.7.

intention. A common intention may, despite defiance in logic, still be found by the court since it would be *unconscionable* not to. However, the vagueness of unconscionability increases judicial uncertainty, and further complicates a field which is already perplexing enough.

Imputation in CICTs

In the event that a common intention cannot be found, imputation is often construed as a 'fail-safe' mechanism. Imputation of a common intention is used in the quantification stage of applying CICTs and resulting trusts.[10] This point is clarified after academic uproar following *Stack v Dowden*, since the imputation mechanism was introduced but its applicability was not qualified by the House of Lords. Imputation was clarified to be used only if a common intention could not be found. Its discussion can also be found in earlier cases concerning the family home, such as *Oxley v Hiscock.*[11] Differing from *Stack v Dowden*, it is a sole legal ownership case, where the claimant contributed to 28 per cent of the property but claimed a 50 per cent interest. The court had settled on 40 per cent, with express discussion of the possibility of imputing a common intention to the parties, in the case where neither adduced any evidence in evidence of one.[12] This is done with reference to the 'whole course of dealing between them [the claimant and defendant] in relation to the property', examples of which include mortgage contributions, council tax payments, and housekeeping.[13]

Imputation is an exercise of the court's jurisdiction. It runs the risk of being a proxy of the court's moral compass. Serious issues include (1) the impossibility of detrimentally relying on an intention that has been imputed by the court (since the claimant cannot rely on something that did not exist prior to court proceedings); and (2) unprincipled judicial discretion, which further obfuscates this area of law. Even after judicial clarification in *Jones v Kernott*, imputation remains controversial since it lacks a sound, algorithmic doctrinal basis. Therefore, it can be viewed that imputation is a problem stemming directly from the need to establish a common intention.

Detrimental Reliance in CICTs

Detrimental reliance is another problem faced using CICTs in the family home context. Prominent feminist academics have argued that, traditionally, women have been discriminated against by the courts; a prominent example is *Burns v Burns.*[14] Therein, Mrs Burns contributed substantially, not only to the family, but to the upkeep of the family home. For instance, she purchased 'certain fixtures and fittings for the house', and 'certain consumer durable goods' which effectively enhanced the value of the property. However, such contributions are deemed as 'housekeeping expenditure.' Therefore, this was classified separately from any financial contribution to the *acquisition* of the property. No reference had been made to how such contributions would *eventually* – albeit indirectly – contribute to such acquisition, e.g., the defendant's expenditure accordingly decreases, increasing his chances of paying off the mortgage. This sparks the acrimonious proposition that gender discrimination is at play: the

[10] *Jones v Kernott* [2011] UKSC 53.
[11] *Oxley v Hiscock* [2004] EWCA Civ 546.
[12] ibid., [65]-[66], [69].
[13] ibid., [69].
[14] [1984] Ch 317.

claimant's contributions are undermined and construed as 'mere evidence of her being a good wife' simply because housekeeping is traditionally viewed as a strictly feminine responsibility; fulfilment of this responsibility does not warrant the 'extra prize' of acquiring a proprietary interest. In colloquial speech, she merely did what her *pay grade required her to do*. This is even more clearly illustrated by *Eves v Eves*[15], where the claimant 'impressed' Lord Denning by having allegedly picked up a sledgehammer and exceeded traditional gender roles by doing 'the men's work' of construction. Developing this idea further, Wong argues that this reek of blatant sexism, since women are assigned specific gender roles and the mere performance of those roles, regardless of how tough the work is, would not amount to detrimental reliance.[16] This statement is praised for its generality, i.e., it is not confined to the familial scenario. Reliance could only be 'detrimental' if the claimant went 'beyond' the fulfilment of such responsibilities, e.g., what a normal female would usually act and behave in a patriarchal society. Only then did the claimant sustain a loss which would warrant judicial reparation (granting a share in the property) as the claimant exceeded her gender-related societal expectations. Again, in colloquial terms, in this case, she 'exceeded her pay grade.'

The recent case of *O'Neil v Holland*[17] may demonstrate the court's acknowledgement of societal emphasis on gender equality since detrimental reliance is more generously interpreted. In this case, even a passive yet drastic change of material circumstances can be interpreted as detrimental reliance. For instance, a female claimant no longer needs to demonstrate her aptitude in carpentry (like in *Eves v Eves*). If the undertaking of a joint venture (regarding the family home) with the defendant entails a material change of circumstances to the claimant which amounts to a substantial (i.e., more than minimal) detriment, this already satisfies the detrimental reliance limb. However, this reveals again the key problem: the parameters of detrimental reliance depend on courtroom justice, instead of long-standing doctrines. The courts are offered broad discretion regarding its definition and are therefore capable of creating discrepancies in different cases. Courtroom adjudication may amount to overt interference towards what the parties deserve.

The Canadian Solution - Contribution-based Approaches

Considering the aforementioned issues, a legal solution is urgently required. This solution must depart from the traditional paradigm of common intention constructive trusts. The Canadian solution is preferred since it has departed from its traditional resulting trust analysis.[18] It introduces constructive trust, which is remedial in nature, hence not requiring a common intention. The remedial constructive trust is based on unjust enrichment. The original resulting trust analysis was similar to the English model, where a common intention was paramount. This method had become increasingly untenable, especially after the controversial case of

[15] ibid, n.6.
[16] Simone Wong, 'The Iniquity of Equity: A Home-Sharer's Tale' [2008] Singapore Journal of Legal Studies, 330.
[17] [2020] EWCA Civ 1583.
[18] Prior to *Pettkus v Becker* (see reference 12), the Canadian courts utilise a resulting trust analysis which resembles the English common intention constructive trust. Although known by different names, they are of the same essence, whereby a common intention is required to be proven by the claimant regarding the allocation of beneficial interest of the property.

Murdoch v Murdoch.[19] In this case, an abused ranch wife was not given a beneficial interest in the property since (1) there was no express, common intention between the two parties; and (2) the detrimental reliance suffered by the wife was deemed 'the same' as 'any' ranch wife. In short, the Supreme Court of Canada fell into the possible problems which would normally arise from CICTs as discussed earlier. Laskin JJ gave a strong dissent. Denouncing the approach adopted by the majority (4-1), he opened up the possibility of using a remedial constructive trust based on unjust enrichment:

> *'In making a substantial contribution of physical labour, as well as a financial contribution, to the acquisition of the successive properties culminating in the acquisition of the Brockway land, the wife had established a right to an interest which it would be inequitable to deny and which, if denied, would result in the unjust enrichment of her husband. Denial would equate her strenuous labours with mere housekeeping chores which, as has been held (see Kowalczuk v. Kowalczuk, [1973] 2 All E.R. 1042), will not per se support a constructive trust. Moreover, the evidence in the present case was consistent with a pooling of effort by the spouses to establish themselves in a ranch operation.'*

From this analysis, the word 'inequitable' is used, in reflection of the roots of the remedial constructive trust, whereby the courts bring about, through judicial operation, changes in property title in reflection of what justice requires to be done. This analysis is not perfect, however. While Laskin J supports the idea of preventing the unjust enrichment of the defendant, he also drew a line of distinction between 'strenuous labours' and 'mere housekeeping chores.' Per my earlier analysis, this should be rejected due to its sexist undertones. It is also not reflective of the current law in England and Wales, since the performance of housekeeping chores and maintaining the care of any dependent children would constitute detrimental reliance.[20] However, Laskin JJ would be excused since his opinion is the primordial form of a novel application of the unjust enrichment doctrine which would ultimately improve distributive justice. Subsequent cases also saw its continual refinement.

This approach is concretised in judicial practice in *Pettkus v Becker.*[21] This landmark decision concerned a sole legal ownership case, where the claimant helped the defendant to build a bee-keeping business and purchase multiple farms as well as properties in both Quebec and Ontario. The requirements for unjust enrichment had been met: (1) the defendant gained a significant monetary benefit and was enriched by the services, (2) at the expense of the claimant, and (3) the enrichment was unjust because of unconscionability. After all, the joint venture between the two parties collapsed. The contributions rendered by the claimant in service of the joint venture should be reimbursed, since there is no evidence that the claimant's contributions are rendered for the defendant's full disposal. The same test is used in English law. Referring to Dickson J:

[19] *Murdoch v Murdoch* [1975] 1 S.C.R. 423.
[20] *Gissing v Gissing* [1970] UKHL 3.
[21] *Petttkus v Becker* [1980] 2 SCR 834.

> '*Where one person in a relationship tantamount to spousal, prejudiced herself in reasonable expectation of receiving an interest in property and the other in the relationship freely accepted benefits conferred by the first person in circumstances he knew or ought to have known of that expectation, it would be unjust to allow the recipient of the benefit to retain it.*'

Canadian common law has therefore recognised a contribution-based approach based on unjust enrichment and remedial constructive trusts.

In terms of trust law, the UK Supreme Court rejected the use of remedial constructive trusts repeatedly in this jurisdiction.[22] Lord Neuberger also delivered a lecture denouncing the use of the remedial constructive trust.[23] He did not notice the irony that by rejecting the remedial constructive trust, and the unfettered judicial discretion which comes with it, he subscribes to the idea of principled judicial discretion. This is at odds with what is happening within CICTs.

It is arguable that remedial constructive trusts should be introduced in England and Wales, in a principled way. Unjust enrichment is a well-established doctrine as compared to unfettered judicial discretion. Certain legal principles, including total failure of consideration, are rooted in the common law. The worries enunciated by the dissenting judges in *Pettkus v Becker* can also be allayed if the judicial discretion is fettered by the need to be exercised with reference to traditional doctrines. However, it is highly unlikely since the Supreme Court decision was announced not long ago – changing a legal position tracing back to *Westdeutsche* might lead to unwelcome legal unpredictability and uncertainty.

Nature of the Enrichment

An important takeaway, prima facie, is that the Canadian approach has shifted its attention from the need for a common intention between the parties, towards the claimant's contributions. Since we are dealing with contributions, which would depend on objective evidence, there is less room for judicial discretion. This potentially improves judicial certainty and promotes distributive justice. However, some may suggest that the adoption of unjust enrichment might cause additional issues, owing to the subjective element of the 'enrichment' limb.[24] This is based on contemporary reality. Although such claims might often involve financial expenditure contributed by the claimant in the maintenance of the home, the claimant is also likely to ask for recompense (in the form of proprietary interest pertinent to the family home) pursuant to their services which improve the state of the property, e.g., gardening, assembling furniture, and rearing of any children. Services operate differently than financial contributions in unjust enrichment. While the latter have a fixed, uniform value, the same cannot be stated for the former. Services are valued differently, according to the needs and personal characteristics of the recipient. Subjective devaluation can take place, in which the

[22] *FHR European Ventures LLP v Cedar Capital Partners LLP* [2014] UKSC 45; *Westdeutsche Landesbank Girozentrale v Islington LBC* [1996] UKHL 12 (Lord Browne-Wilkinson).
[23] Lord Neuberger, "The Remedial Constructive Trust — Fact or Fiction" (Speech given at the Banking Services and Finance Law Association Conference, Queenstown in August 2014) < https://na.eventscloud.com/file_uploads/ebf86fe315a881f65326d99209ad69af_P1-LordNeuberger.pdf> accessed on 20 March 2021.
[24] *Bank of Cyprus v Menelaou* [2015] UKSC 66, [20].

recipient claims that the service renders them no benefit at all, so that the service is not an enrichment in the first place. As suggested by Gardner,[25] subjective devaluation would not be possible if either (1) the recipient requested for that service to be rendered (this *may* be relevant in the family home context, but would be fact-specific); or (2) it would be unconscientious for the recipient to do so, provided that (a) the provider does so expecting recompense (hence not gratuitous), and (b) the recipient is cognisant of (a). However, this would lead to additional problems – what does the recipient subjectively know? Furthermore, does the provider genuinely expect recompense? The latter question is particularly problematic, because it could be suggested that partners in a relationship may be driven to contribute by *emotional rewards*, not proprietary rewards.

Affronted by this dilemma, slight doctrinal modifications can be made. The New Zealand approach can be adopted. According to Cooke P. in *Gillies v Keogh*,[26] the correct question is whether a reasonable person in the claimant's position would have expected an interest. This is a purely objective test and is fact specific. It would depend on the nature and development of the relationship in question. This approach better protects the claimant's interest. In cases where the claimant has moved into the defendant's home, it is no longer sufficient for the claimant's interest to be denied where the defendant alleges that the mere fact that the couple opted *not* to purchase a separate property but agreed that the claimant move to the defendant's home, induces the defendant to believe that the claimant rendered them services for reasons which are *not* related to financial recompense pertinent to the family home.

It must be noted that the family context is different from the commercial context. Trust and collaboration are the pillars of a romantic and familial bond, as emphasised by Gardner.[27] Moreover, cohabiting implies the creation of a joint venture involving the family home, signifying more advanced financial integration between the parties. Objectively speaking, in the general case, for a relationship to have progressed to the stage where two parties cohabit, the degree of trust between the parties and confidence in the joint venture is so great that a mere assertion from the defendant that the claimant will not get an interest in the family home is not sufficient. This would be incongruent with the motive of cohabiting in the first place, when viewed objectively. It thus follows that: even if the defendant told the claimant explicitly that they would not acquire an interest in the family home, *objectively viewed*, it is still reasonable for the claimant to expect recompense from the services they rendered, given the progress of the relationship and degree of financial integration normally expected from a cohabiting relationship. Moreover, also deduced from the very stage of the relationship (and financial integration) cohabitation implies, the recompense expected is likely to include both emotional, and proprietary rewards. The venture is no longer limited to personal, emotional gratification which is usually more prominent at the earlier stages of a romantic association. It also involves the financial side of affairs.

Of course, in line with distributive justice, there should be ways in which this objective presumption can be rebutted, so that the defendant can subjectively devalue the services rendered. Evidence can be adduced to prove that cohabitation is the result of circumstances

[25] ibid., n 8, [284].
[26] *Gillies v Keogh* [1989] 2 N.Z.L.R. 327.
[27] ibid., n 8, [283].

which are independent of the progress of the relationship, e.g., the claimant is evicted from her home where the defendant accommodates them out of mere generosity, or that the defendant has a medical condition like epilepsy which requires the claimant's supervision. Any services rendered by the claimant will be conducted against the backdrop of a less *mature* relationship. Hence, they are most likely rendered because of affection and/or necessity, rather than investment in the joint venture. It would therefore not necessarily be unconscionable for the defendant to subjectively devalue the services rendered in relation to the family home.

Proprietary Remedies in Unjust Enrichment

There has been considerable debate as to the *form* of the remedy granted. Although unjust enrichment explains why the claimant should be reimbursed for their efforts, it is not an adequate explanation as to why the reimbursement takes the form of a *proprietary* interest in the family home. A proprietary remedy is arguably more valuable than a personal remedy, since (1) the claimant can benefit from increases in the market value of the property; and (2) the claimant can gain priority over uninsured creditors in the event of the defendant's bankruptcy. Moreover, property-based personal remedies can also satisfy the issue of labelling, where the remedies acknowledge the former existence of a joint venture concerning the family home.

There are certain recognised doctrines in English law which confer property rights owing to unjust enrichment. It is indeed not spontaneous judicial innovation, nor would it require extensive doctrinal restructuring. One prominent example is subrogation. Adopting a simplified view of the doctrine, subrogation occurs when C gives money to A to pay off A's debts to B. In a mortgage scenario, Bank C, upon contributing to the satisfaction of Bank B's debts by Individual A, has a charge over A's property. An example is the family home case of *Graham-York v York.*[28] Although doubted by Australian case law[29], an unjust enrichment analysis as the doctrinal underpinning of subrogation is recognised in English courts.[30] The reasons behind the grant of a proprietary remedy in this case warrant serious consideration. This further protects the rights of the lender, since they can be prioritised over other uninsured lenders upon the borrower's insolvency. Insolvency is far from a remote possibility in this case, since the very act of borrowing money indicates an intrinsic deficiency in financial management and business acumen. Moreover, with regard to social policy, this supplies loan providers with another attraction to abstain from increasing current interest rates, which might affect the local economy.

Burrows is a strong proponent of granting proprietary remedies under unjust enrichment[31], subject to certain exceptions. His theory stresses that the property that the defendant has gained from the claimant is 'at the claimant's expense.' In equity, a trust is established in the claimant's favour where the defendant is under the obligation to return the property. To commence with, a proprietary remedy is more efficacious in protecting the claimant's rights. Greater protection is warranted in this case, since, if it were not for the

[28] *Graham-York v York* [2015] EWCA Civ 72.
[29] *Bofinger v Kingsway Group Ltd* [2009] HCA 44.
[30] *Filby v Mortgage Express Ltd (No. 2)* [2004] EWCA Civ 759, [62] (Lord Justice May).
[31] Andrew Burrows, 'Proprietary restitution: unmasking unjust enrichment' [2001] 117 LQR 2, 417.

claimant's contribution to the joint venture, it would be doubtful whether the defendant themselves could acquire the property successfully in the first place. Moreover, the money paid by the claimant to the defendant contributes to the joint family venture and increases the value of the property (or properties) concerned. Any monies used otherwise upon the collapse of the joint venture (i.e., the breakdown of the relationship), such as for the defendant's own benefit, should be held on trust for the claimant.

In addition, to be eligible for a proprietary remedy, the claimant needs to act as if he/she is a secured creditor who has not taken the risk of the defendant's insolvency. This view is endorsed by Lionel Smith.[32] He reasons that in a scenario resembling *Foskett v McKeown*[33], where the stolen trust monies are disposed of, the trust subsists regardless of the acts the trustees committed. The trustee is still bound by the trust and has to account for the beneficiaries of said trust. However, if the trust monies were handed to a third party, it would be aberrant to suggest that the original trust could bind a third party, requiring the third party to hand over the trust monies to the beneficiaries, provided that the third party acted in good faith with no knowledge of the trust. Beneficiaries cannot sue third parties regarding infringement of property rights.[34] In the family home situation, the claimant has not reasonably taken the risk of the defendant's insolvency by contributing to the joint venture. In many cases, the claimant is unlikely to be aware of the fine details of the financial situation of the defendant, since the nature of the relationship remains one of emotion and affection, rather than strictly commercial. It can be further suggested that it is unfair to bar the claimant from recovering any unspent contributions simply because of a presumption of the claimant's knowing receipt of the defendant's financial situation. Despite having known about the defendant's insolvency, actual or imminent, the claimant cannot be reasonably taken to have the obligation to be acquainted with the legal consequences of receiving such knowledge. Suggesting that claimants should be bound, deviates from reality where couples are, as aforementioned, not always apt to discuss long-term financial arrangements in staggering detail.

Quistclose Trusts and Resulting Trusts

Proprietary remedies under unjust enrichment remain controversial. As a result, some commentators such as Lord Millett and Graham Virgo have suggested trust law as the doctrinal basis of tracing proceeds, *Quistclose* trusts are suggested as the primary modality. Lord Wilberforce stated in the *Quistclose* case:[35]

> *'When the money is advanced, the lender acquires an equitable right to see that it is applied for the primary designated purpose (*see Re Rogers (U.S.) where both Lindley L.J. and Kay L.J. explicitly recognised this*): when the purpose has been carried out (i.e. the debt paid) the lender has his remedy against the borrower in debt: if the primary purpose cannot be carried out, the question arises if a secondary purpose (i.e. repayment to the lender) has been agreed, expressly or by implication: if it has, the remedies of equity*

[32] Lionel Smith, 'Unravelling proprietary restitution' [2004] 40 Canadian Business Law Journal, 317.
[33] *Foskett v McKeown* [2000] UKHL 29.
[34] c.f. *Shell UK v Total UK* [2010] EWCA Civ 180.
[35] *Barclays Bank Ltd v Quistclose Investments Ltd* [1968] UKHL 4, 656B-D.

> *may be invoked to give effect to it, if it has not (and the money is intended to fall within the general fund of the debtor's assets) then there is the appropriate remedy for recovery of a loan. [emphases mine]''*

In short: A (the claimant) gives B (the defendant) £100 for a specified purpose - any use of the money in service of a purpose other than the one proposed by A, should be held on trust for A. B is therefore accountable to A as to how the money is used. Granting a proprietary remedy in this case more effectively protects the rights of the lender. Moreover, doctrinally speaking, B acts as the fiduciary of A's assets, justifying the imposition of a trust. The transfer of legal title of the property from A to B has never been unconditional – it is conditional upon the satisfactory performance of pre-agreed objectives. Arguing that A is only eligible for a personal remedy would deviate from sound doctrinal reason.

In the family home context: If A gives a sum to B for the purpose of the joint family venture. A does not intend for B to use the money for their own disposal. A intends for the monies to be used for the *realisation* or *smooth operation* of the joint venture. Any unspent monies, or monies used for a purpose alternative to the joint venture, must be held on trust for A. Therefore, upon relationship breakdown, where the joint venture accommodates contributions by both A and B, the rules of equity in tracing allows A to reclaim the exact amount A contributed to B during the joint venture.

Two criticisms should be addressed: (1) *Quistclose* trusts have traditionally been applied in commercial cases. Understandably, most cases concern monies instead of services. It would be an unjustifiable doctrinal stretch to apply this doctrine in the domestic realm; and (2) Domestic contributions are not always well-demarcated. The joint venture in question may expand beyond the confines of the family home and include chattels, such as the family car. *Quistclose* trusts should be used in situations where financial contributions are much clearer.

In response to (1), *Quistclose* trusts have never been deployed in the family home context, not because of their inapplicability, but because of their doctrinal origins and the novelty of this proposition. Heed that the index *Quistclose* case is commercial in nature and concerns a debtor-debtee relationship. However, there is no sensible reason why the doctrine (1) cannot be applied to the family context due to doctrinal reasonability; and (2) cannot be expanded to include services. Housekeeping services can naturally be quantified by adducing evidence of expenditure. Such expenditure is directly linked to the acquisition of the property, since, for instance, it signifies the sums saved by the defendant to pay off the mortgage attached to the family home. As for services pertinent to the improvement of the home environment, although they may vary in value, they can be quantified by comparing market prices. To illustrate, we have: (1) the value of the family home at the point of separation (since then, it would become unconscionable for the defendant to keep the contributions, for the joint venture would have come to an end); and (2) the value of a nearly identical/similar property without the household modifications in question. Each case would turn on its particular facts. Understandably, litigation proceedings would be rather long due to the sheer amount of evidence incurred for each limb. However, even under the current CICT doctrine, extensive evidence is often adduced to rebut the presumption of equivalent shares in the property.

In response to (2), the basis of the criticism is the lack of clear demarcation of purpose regarding the claimant's financial contributions. This criticism originates from a commercial

mindset which is not necessarily applicable to a domestic household. Indeed, there are two alternative ways of interpretation. Firstly, the chattel recipient of the contribution might be deemed judicially as an integral part of the family home.[36] This can be exemplified by fixing a shed attached loosely to the main body of the family home. Secondly, owing to the intricate relationships between different strands of expenditure in the family home, the joint venture is better interpreted. Overemphasis on the minutiae is not encouraged. For instance, the claimant might have financially contributed to the repair of the family car. However, it does not automatically follow that the contribution is not related to the family home. The financial contribution is directed at the joint venture mentioned above. Financial integration of both parties is so strong that, at this stage, the claimant's financial contribution can be said to save the defendant's expenditure effectively enough that the defendant could affect beneficial improvements on the family home or keep up with mortgage payments. Therefore, although the claimant's financial contributions may be directed at different components, they are conjoined to drive the joint venture forward. The imposition of a *Quistclose* trust is thus justified.

Due to the persistent rejection of the resulting trust analysis by English courts, with Lady Hale announcing that the law has indeed moved on,[37] it would be paradoxical and create substantial judicial uncertainty should we return to the 'resulting trust era.' However, the earlier rejection of the resulting trust was based on modern societal attitudes towards the changing roles of women and relationships, as well as the recognition of the family home as part of a joint venture. There would be no paradox if a return to this analysis signified the reinforcement of these ideals. Moreover, *Quistclose* trusts can be considered as a species of resulting trust. This was first expressed by Lord Millet in *Twinsectra v Yardley*, where he declared that *Quistclose* trusts were 'illusory.'[38]

However, certain ambiguities in the resulting trust analysis should be noted. This can affect the feasibility of this approach in legal practice. There is still considerable academic debate as to what the 'presumption' of underlying resulting trusts truly is: (1) a positive presumption that A has given the property to B so that B can hold it on trust in favour of A; alternatively, (2) a negative presumption that A has not intended for the property to be a gift to B? The difference is minimal yet significant, vociferously argued by Chambers.[39] It has also been authoritatively argued that judges use the two formulations interchangeably. This is significant in the family home context as the formulations differ in the methods required to discharge the evidential burden. Should presumption (1) be adopted, to disprove it, B could adduce evidence of A's reluctance to incur any taxes.[40] A would therefore prefer renouncing all forms of interest associated with the transferred property. If presumption (2) were adopted, such evidence would fail to discharge the evidential burden, since the unwillingness to incur taxes would not automatically entail that A has intended to give B a gift. Using either

[36] *Elitestone Ltd v Morris* [1997] 1 WLR 687.
[37] ibid., n 1, [60].
[38] *Twinsectra Ltd v Yardley* [2002] UKHL 12.
[39] John Mee, 'Presumed Resulting Trusts, Intention and Declaration' [2014] 73 CLJ 1, 102-104; Robert Chambers, *Resulting Trusts* (Clarendon Press, 1997), 19.
[40] *Re Vandervell Trustees Ltd (No 2)* [1974] EWCA Civ 7.

presumption, A's beneficial interest in the property (resultant from A's contributions) can still be secured and can be reclaimed upon the breakdown of the joint venture.

In the area of trust law, some critics are concerned with the labelling effect. Some may suggest that the use of a common intention constructive trust carries a more profound meaning than merely asset division. It is a cultural signifier whereby the law recognises the intrinsic value of the family home and real property. Plots of land are no longer merely seen as assets. The family home has been suggested as a 'special type of property'[41]. It is a basic unit of society, which gives character and essence to a community. In fact, current academic opinion argues that traditional land law doctrines such as overreaching, emphasised, and consolidated in statutory legislation against a post-war landscape, overly focus on the idea that real property is merely an asset and should be rejected to better reflect its contemporary cultural value.[42]While this argument undoubtedly has traction, it should ultimately be rejected. In a case of asset division after relationship breakdown, with additional consideration to social policy, the focus should be on the party who is not given full credit for their contributions during the joint family venture. The want for justice arises when the interests of the individual are not sufficiently safeguarded. The recognition of the joint family venture, or indeed the changing social roles of the family home and real property, does not require the use of a controversial trust device which operates at the expense of the claimant. Besides, using proprietary remedies as mentioned above already recognises the existence of the joint family venture, without the need to establish a common intention.

Conclusion

To resolve the ills of the common intention constructive trust in the context of asset division in common law partners, instead of forcefully finding a common intention or imputing one, greater focus should be placed on the contributions made by the claimant towards the joint venture – in this case, the hands of the defendant. There is no need to adopt the revolutionary move of introducing remedial constructive trusts to English shores. Utilising the concept of unjust enrichment and conventional trust law principles, proprietary remedies can be justified.

[41] Lorna Fox O'Mahony, 'The Idea of Home in Law' [2005] 2 Home Cultures 1, 40.
[42] Dermot Cahill and John Owen, 'Overreaching- Getting the Right Balance' [2017] 81 Conveyancer and Property Lawyer 1, 35.

How compatible is international criminal justice within the African context?

Edward Armitage

Introduction

The compatibility between international criminal justice and the African continent is greatly contested. International criminal justice concerns the accountability of individuals who have committed the most serious crimes, such as, *inter alia*, genocide, crimes against humanity, war crimes and the crime of aggression. The international community has utilised several types of justice mechanisms in the African continent for fighting against impunity, including the International Criminal Tribunal for Rwanda (ICTR); the Special Court of Sierra Leone (SCSL); and, the International Criminal Court (ICC), a permanent court situated in The Hague. Acknowledging Africa's debilitating history with colonialism will be critical to understanding how international law developed and how the neo-colonial practises of the ICC today severely curtail the ability of international criminal justice and Africa to ever be considered compatible with one another. This essay will then go on to examine the problematic formulation of victims and perpetrators on the international stage; how international criminal justice silences male victims of sexual violence; question whether the ICC destabilises or promotes the peace process in post-conflict regions; and lastly, examine whether international criminal justice is too far removed from those in Africa who seek its benefit the most. The arguments presented in this essay will ultimately conclude that the compatibility of international criminal justice and the African continent is severely limited.

Neo-colonial Context

It is imperative to contextualise the suitability of international criminal justice and Africa within the parameters of the latter's extremely coercive and devastating colonial past. The formal colonial period (1880-1940) was a European conquest that sought to under develop the African continent economically, politically and culturally.[43] During the formal partition, as Joireman posits, 'colonial metropoles established their own systems of law and dispute resolution' and pushed pre-existing legal practises to the peripheries, brandishing them as 'primitive' or 'for natives only.'[44] The actions of European colonists lay the foundations for creating a hierarchy and superiority for judicial practises, rendering localised mechanisms of justice incompatible and insufficient to the standards of international criminal justice. Furthermore, it is significant to acknowledge the evolution of international law itself and how colonial influence was a key facet to its evolvement, having the explicit purpose of 'suppressing the Third World.'[45] It is difficult to conceptualise how international criminal justice could ever be considered to be well suited in the African context when the foundations of international

[43] Walter Rodney, *How Europe Underdeveloped Africa* (Pambazuka Press, 2012).
[44] Sandra Joireman, 'Inherited legal systems and effective rule of law: Africa and the colonial legacy' [2001] 39 The Journal of Modern African Studies 4, 571.
[45] Antony Anghie, 'The evolution of international law: Colonial and postcolonial realities' [2005] 27 Third world quarterly 5, 748.

law are inextricably linked to colonialism and imperialism, which only seek to impose power and control over the colonised.

Whilst the cessation of formal colonialism may warrant fresh sentiments of independence for postcolonial states, Western countries introduced a more subtle regime of power and control known as neo-colonialism. Nkrumah defines neo-colonialism as a 'State which is…in theory, independent and has all the outward trappings of international sovereignty' but '…[I]n reality its economic system and thus its political policy is directed from outside.'[46] The ICC has attracted extensive criticism as being a neo-colonial power due to its selectivity in investigations which, up until recently, has focused wholly on the African continent. Ten out of the fourteen investigations opened by the ICC are in Africa, with the remaining four being conducted in the Global South. The power disparity between African nations and Western states is further exacerbated by the ability of the United Nations Security Council (UNSC) to refer situations to the ICC under Chapter VII of the UN Charter. This means that powerful governments who have not ratified the Rome Statute within their own domestic legal systems are still able to influence the ICC without being held accountable themselves. The UN's referral of the Darfur situation to the ICC was welcomed by the US, despite vehemently opposing the ICC and labelling it as 'Kangaroo Court.'[47] The readiness of the ICC to arrest the acting Head of State of Sudan but not to do the same for the US and UK governments with their unjustifiable invasion of Iraq[48] highlights the hierarchy of crimes worthy of prosecution in the eyes of international criminal justice. Therefore, the neo-colonial practises of the ICC result in more powerful states enjoying a degree of impunity whilst less powerful countries are not afforded such luxuries. Whilst this should not exempt prosecutions of crimes that have been committed in African countries, exclusively investigating human rights abuses on the continent fuels the divisive stereotypes created by the colonial mission of the 'primitive' and 'savage' Africa.[49] However, through successfully portraying individuals from the developing world as 'uncivilised', more powerful nations are able to justify intervention.[50] The asymmetric power and control which Western states exert over poorer nations through the ICC highlights the severe incompatibility of international criminal justice within the African context. It is questionable whether international criminal justice was ever intended to be well suited for Africa due to its inseparable ties to colonialism and the protection it affords to powerful countries. At the very least, the incessant concentration on Africa has led to a significant breakdown of diplomatic relations between the African Union (AU) and the ICC, making the prospect of administering international criminal justice even more difficult.

Victim- Perpetrator Dichotomy

[46] Kwame Nkrumah, *Neo-colonialism: the Last Stage of Imperialism* (Panaf Ltd 1974).
[47] Robert Cryer, et al., *An Introduction to International Criminal Law and Procedure* (3rd edn, Cambridge 2014), 164. *See also* Diane Amann, 'The United States of America and the international criminal court' [2002] 50 American Journal of Comparative Law, 385.
[48] Owen Bowcott, 'ICC abandons inquiry into alleged British war crimes in Iraq' (*The Guardian*, 2020). <https://www.theguardian.com/uk-news/2020/dec/09/icc-abandons-inquiry-into-alleged-british-war-crimes-in-iraq> accessed 18 February 2022.
[49] Joseph Conrad, *Heart of Darkness* (Penguin Classics, 1994).
[50] Élise Féron, 'Wartime Sexual Violence Against Men: Why So Obvious?' [2017] 4 European Review of International Studies, 72.

A criminal act produces an individual who has been victimised – the 'victim' – and also an individual who has committed the crime – the 'perpetrator.' Whilst this may naturally occur, a problem arises when there becomes an archetypal definition of the 'ideal victim.' In the field of criminology, Nils Christie defines the ideal victim as 'a person or category of individuals who when hit by crime most readily is given the complete and legitimate status of being a victim.'[51] Christie further attributes several characteristics that form the ideal victim: (S)he is i) weak; ii) carrying out a respectable project; and iii) not blameworthy. (S)he would also have to be victimised by iv) an offender who was big and bad; and also, v) unknown (ibid.).[52] Viewing the 'ideal victim' on the international stage, these characteristics can be supported by Erica Bouris who argues that transitional justice tends to use simplistic categories of 'victim' and 'perpetrator', and states that this simplicity results with a less recognised view amongst the international community of a complicated victimised individual.[53] Bouris discusses the 'ideal victim' is one that is associated with characteristics such as 'innocence', 'purity', 'lack of responsibility' and 'moral superiority', whilst perpetrators are associated with 'evil' and 'guilt.'[54] From a gendered perspective, Ní Aoláin posits how international criminal justice largely negates women's political agency during war as they are seen as 'homogeneously powerless or as implicit victims, thereby excluding…women as benefactors of oppression or the perpetrators of catastrophes.'[55] The ICTR, for example, was primarily focused on holding men to account as they were more generally associated with the characteristics of a perpetrator.[56] It has been stated that when there are clear parameters around 'victim' and 'perpetrator', it is much easier to administer humanitarian aid and retributive justice.[57]

The African continent has witnessed a plethora of violent conflicts in the last fifty years, and due to the type of the crimes that have been committed and the individuals who have committed them, these conflicts have challenged the way international criminal justice conceptualises victims and perpetrators. These questions arise in situations where individuals have been victimised but have also operated in a capacity to become perpetrators; individuals Bouris refers to as 'complex political victims.'[58] Child soldiers are an example where the distinction between victims and perpetrators becomes very ambiguous. Under multiple acts of international legislation (see The Rome Statute; The African Charter in the Rights and Welfare of the Child), it is illegal to recruit children in armed conflict below a certain age. Therefore, child soldiers are victimised in their abduction into armed groups, but then the process of indoctrination, which usually involves horrific and dehumanising methods,[59] combined with their actions in the armed group, means children also cross the Rubicon to become perpetrators. The dichotomy between victims and perpetrators means it is difficult for the international community to know how to perceive and/or address individuals who hold multiple identities.

[51] Nils Christie, *The Ideal Victim* (Macmillan, 1986), 18.
[52] ibid., 19.
[53] Erica Bouris, *Complex political victims*, (Kumarian Press, 2007), 20.
[54] ibid.
[55] Fionnuala Ní Aoláin et al., *On the Frontlines: Gender, War, and the Post-Conflict Process* (OUP, 2011), 42.
[56] ibid.
[57] Liisa Malkki, *Speechless Emissaries: Refugees, Humanitarianism and Dehistoricization*. (Cultural Anthropology, 1996).
[58] Bouris (n 12).
[59] Alicinda Honwana, *Child Soldiers in Africa* (Philadelphia, 2011).

However, with regards to child soldiers, the international community has swept the issue of complex political victims under the rug, and by virtue of their adolescence, child soldiers are viewed wholly as innocent victims.[60]

Therefore, the case of Dominic Ongwen has forced the ICC to formally address such individuals. Ongwen was abducted at the age of nine by the Lord's Resistance Army (LRA) but worked his way up to achieve the rank of Brigade Commander. Thus, he is both a victim due to his abduction as a child and also a perpetrator for the crimes he committed as an adult. The rigid definition of the 'ideal victim' prohibits Ongwen to be seen in such a capacity, as his actions as a perpetrator are neither 'innocent' nor 'pure' but rather 'evil' and 'wicked.' Thus, by virtue of his guilty verdict, Ongwen has only been viewed in his capacity as a criminal, even though if he had not been abducted at the age of nine, he likely would never have committed the crimes that he did. Ongwen's sentence of twenty-five years' imprisonment has left an unsatisfactory conclusion to the issue of complex political victims. Can we really say that Ongwen has had justice served? A slightly reduced sentence does not seem to adequately acknowledge Ongwen's continued status as a victim and appears to be a tokenistic gesture. There is concern that by failing to recognise complex political victims in justice pursuits after conflict, a new space is created in which 'mass victimisation, particularly genocide' can take place.[61] This means that their exclusion from accessing justice potentially fuels the social construction of the 'Other', and this construction is the first step towards dehumanising a subgroup which can lead to violence.[62] Bouris explains that by 'recognising these perpetrators as victims is quite critical, because if we do not see them as victims, we are unlikely to understand the true horror of [their actions].'[63] Therefore, the rigid and dichotomous definitions of the ideal 'victim' and 'perpetrator' are not only discernibly gendered, but also prevents complex political victims to be seen in their capacity as victims. The continued use of children involved in armed conflict in Africa,[64] and the failure of the ICC to properly address complex political victims, portrays that international criminal justice has been, and continues to be, ill-suited to the African context.

Sexual Violence on Men

Historically, sexual violence against men and boys has been a neglected area on the international stage with the media focusing largely on sexual crimes targeted against women and shielding away from the topic of sexually assaulted men.[65] While international legal statutes have adopted neutral definitions of rape that allow men, women, boys and girls to be victims of rape, other 'modes of sexual violence commonly directed at men and boys…are not

[60] Noëlle Quénivet, 'Does and Should International Law Prohibit the Prosecution of Children for War Crimes?' [2017] 28 European Journal of International Law 2.
[61] Erin Baines, 'Complex political perpetrators: reflections on Dominic Ongwen' [2009] 47 The Journal of Modern African Studies 2, 177.
[62] ibid.
[63] Bouris (n 11), 67.
[64] See Honwana (n 17).
[65] Philipp Schulz, 'The "ethical loneliness" of male sexual violence survivors in Northern Uganda: gendered reflections on silencing' [2018] 20 International Feminist Journal of Politics 4, 583.

explicitly listed.'[66] Positive steps have been made by the ICC whereby, under Articles 7 and 8, the Rome Statute does provide a more comprehensive list of sexual crimes. However, the lack of recognition of sexual violence against males has warranted very few prosecutions in international justice mechanisms.[67] There is limited case law on the matter, with the ICTY contributing the most and the ICTR, ICC, and the SCSL providing some additional supportive jurisprudence.[68] However, what is most deafening are instances when charges were not brought forward against male sexual violence. In the SCSL, even though the Trial Chamber acknowledged sexual violence towards men, the Prosecutor restricted indictments of sexual violence to those only directed against civilian women and girls.[69] The inconsistency of international justice tribunals highlights that there is 'no overarching or coherent prosecutorial policy…on how to approach this form of sexual violence.'[70]

Part of the failure of international criminal justice to address sexual violence against men can be rooted within the hetero-normative and patriarchal values upheld within the African context, whereby male victims to sexual violence suffer immediate effects to their gender identities.[71] These effects are compounded during periods of conflict when 'hyper masculinities' such as strength and aggressiveness are elevated.[72] Consequently, many male survivors are silenced because they feel reluctant to come forward to report their victimisation due to the 'shame, confusion, guilt, fear, and stigma' associated with male sexual violence.[73] Many male survivors have been left by their wives as they have been perceived to have 'lost their manhood' after being raped and thus, their wives did not want to live with 'a fellow woman.'[74] The silencing of male victims is further exacerbated when combined with the reality that many African countries have made homosexuality illegal. For example, the Ugandan Parliament introduced the Anti- Homosexuality Act 2014, prohibiting sexual relations between persons of the same sex, whilst the Ugandan Penal Code omits males from the definition of rape. Thus, male victims find themselves in a state of 'ethical loneliness' which Stauffer defines as the 'isolation one feels when one, as a violated person or as one member of a persecuted group, has been abandoned by humanity, or by those who have power over one's life's possibilities.'[75] The 'ethical loneliness' of male victims is exacerbated by the fact that the personnel involved in international criminal justice such as investigators, prosecutors and judges 'may face challenges in recognising sexual violence.'[76]

Therefore, men and boys who are victims of sexual violence find themselves at a very complex conjuncture. The lack of recognition of male sexual violence by international criminal

[66] Valerie Oosterveld, 'Sexual Violence Directed Against Men and Boys in Armed Conflict or Mass Atrocity: Addressing a Gendered Harm in International Criminal Tribunals' [2014] 10 Journal of International Law and International Relations, 121.
[67] ibid.,110.
[68] ibid. See also *The Prosecutor v. Jean-Pierre Bemba Gombo* ICC-01/05-01/08.
[69] Oosterveld (n 24), 112.
[70] ibid.
[71] Schulz (n 23), 592.
[72] Ní Aoláin (n 13), 49.
[73] Sandesh Sivakumaran, 'Lost in translation: UN responses to sexual violence against men and boys in situations of armed conflict' [2010] 92 International Review of the Red Cross 877, 1288.
[74] Schulz (n 23), 592.
[75] Jill Stauffer, *Ethical Loneliness: The Injustice of Not Being Heard* (Columbia University Press, 2015).
[7676] Oosterveld (n 24), 119.

justice mechanisms and the failure to prosecute sexual crimes against men means that male victims are not heard on the international stage. Moreover, the position of homosexuals in many African countries results with male victims not being able to turn to their domestic jurisdictions either. Thus, failure to address the complexities of male sexual violence, particularly during instances of hypermasculinity and hetero-normative settings, makes international criminal justice ill-equipped and ill-suited to operate within the African context.

Peace versus Justice

Prominent discourse surrounding international criminal justice is whether it enables or destabilises the peace process in post conflict regions. The most notable incident in peace versus justice discussions is the arrest warrant for President Al-Bashir issued by the ICC. Sudan was in a state of fragility when the arrest warrant was issued, which resulted with the AU requesting the UNSC to defer the indictment so peace negotiations could be conducted.[77] When this request was refused, the AU instructed its Member States not to cooperate with the ICC's arrest warrant for Al-Bashir.[78] This incident demonstrated how the ICC's intervention posed a very real risk of destabilising peace in Sudan but also how the lack of cooperation 'created suspicion and lack of trust between the ICC and AU in fighting impunity on the continent.'[79]

The ICC's intervention in the conflict in northern Uganda is another example of how international criminal justice is ill-suited to the African context. After many years of intense fighting in Northern Uganda, the Ugandan government passed the Amnesty Act, which provided amnesty procedures for fighters in the LRA as well as other rebel groups that resisted government forces.[80] However, it was in the midst of immensely fragile peace negotiations in which the ICC, contrary to the Amnesty Act, issued arrest warrants for the most senior LRA commanders to hold to account those 'bearing the greatest responsibility.'[81] Not surprisingly, the reaction to such intervention was met with grave concern among human rights organisations. Allen notes that the arrest warrants 'would practically close… the path to peaceful negotiation… crushing whatever little progress has been made over these years.'[82] This opinion was corroborated by an Amnesty Commission spokesperson who agreed that the arrest warrants would make it very difficult 'for the LRA to stop doing what they are doing.'[83]

The actions of the ICC in these scenarios demonstrate several reasons why international criminal justice has proven to be ill-suited to the African context. First, it exemplifies complete disregard for the impact issuing arrest warrants might have on the progress to ending a very

[77] Kamari Clarke & Sarah- Jane Koulen, 'The Legal Politics of the Article 16 Decision: The International Criminal Court, the UN Security Council and Ontologies of a Contemporary Compromise' [2014] 7 African Journal of Legal Studies 3, 309.
[78] ibid.
[79] ibid.
[80] Tim Allen, *Trial Justice: The International Criminal Court and the Lord's Resistance Army* (Zed Books, 2006), 74.
[81] ICC Press Release, 'ICC - President of Uganda refers situation concerning the Lord's Resistance Army (LRA) to the ICC' (2004) <https://www.icc-cpi.int/Pages/item.aspx?name=president+of+uganda+refers+situation+concerning+the+lord_s+resistance+army+_lra_+to+the+icc> accessed 18 February 2021.
[82] Allen (n 38), 85.
[83] Allen (n 38), 86.

violent war for the Ugandan people. Branch notes that 'the execution of the arrest warrants would require a dramatic intensification of the government's counterinsurgency in order to capture the LRA leaders.'[84] Second, the ICC's intervention highlights their lack of knowledge pertaining to a very complex war where the main victims in this conflict are children. Furthermore, considering that the ICC intervened in Uganda in 2003, it is submitted that the ICC had not learned from its past mistakes when it subsequently intervened in Sudan six years later. The similarities in both scenarios with regards to the fragility of peace negotiations when ICC intervention occurred demonstrates that the goals of international criminal justice have consistently not aligned with the desires of the country in which the violence has occurred. Thus, the arrogance of ICC and their persistent desire to end impunity only exacerbates the incompatibility of international criminal justice and the African continent even further.

Despite the weight of these arguments, Clark questions whether it is appropriate to label the ICC as an obstacle to peace in Africa given that (at the point of issuing the arrest warrants for LRA commanders), 'there has been no real peace in northern Uganda since 1986.'[85] Furthermore, an agreed peace accord between government officials and rebel groups does not necessarily guarantee peace itself due to the fact that 'violence is not switched off like a tap….'[86] However, simply due to there being no previous peace deal in place, it does not negate the fact that arrest warrants issued by the ICC have made the reality even more unattainable. Moreover, whilst it would certainly be naive to assume that violence can be 'switched off like a tap', a peace accord is irrefutably a move in the right direction.

Additionally, Clark contends that whilst the ICC's arrest warrants may temporarily pose a threat to negative peace (i.e., the absence of violence), by indicting perpetrators for their crimes, 'further down the road it may be exactly what is needed to get a stable peace.'[87] Therefore, support for punitive measures echoes sentiments of 'no justice, no peace' and even though prosecutions may disrupt a region briefly, in the long term indictments will be more beneficial for a longer-lasting and more stable society. Unfortunately, however, this strategy does not reflect the reality within the African context where the ICC has achieved convictions. Recent NGO accounts from Mali and the Democratic Republic of the Congo[88] have reported serious acts of violence committed in both countries which call into question the success of criminal convictions in securing a stable peace in post-conflict societies.[89].

It is further problematic to contend that punitive measures promote a longer lasting form of peace in post-conflict regions because it assumes that international criminal justice has best served everyone affected by the conflict. However, the ability of state actors to manipulate

[84] Adam Branch, 'International Justice, Local Injustice' [2004] *Dissent* <https://www.dissentmagazine.org/article/international-justice-local-injustice> accessed 18 February 2022.
[85] Janine Natalya Clark, 'Peace, Justice and the International Criminal Court' [2011] 9 Journal of International Criminal Justice 3, 542.
[86] Roger Mac Ginty, *No War, No Peace: The Rejuvenation of Stalled Peace Processes and Peace Accords* (Palgrave Macmillan, 2006).
[87] Clark (n 43), 542.
[88] The ICC has obtained only four convictions. One from the situation in Mali and the remaining three from the situation in the DRC.
[89] United Nations 'Democratic Republic of the Congo' (The Office of the Special Representative of the Secretary-General on Sexual Violence in Conflict, 2021) <https://www.un.org/sexualviolenceinconflict/countries/democratic-republic-of-the-congo/> accessed 18 February 2022.

international criminal justice mechanisms to their advantage results in many individuals going unpunished for their actions because they possess power on the political and military level. In the ICTR, for example, no prosecutions were brought against leaders of the Rwandan Patriotic Front (RPF), despite findings from human rights organisations that accused the RPF of murdering thousands of unarmed civilians.[90] Similarly, self-referrals to the ICC have proven favourable to state leaders who trade off state cooperation for impunity. Criticisms have been made pertaining to the self-referrals of Uganda, the DRC, Mali and the Central African Republic, that state leaders have 'been able to use the ICC for their own political and legal gains.'[91] Nouwen affirms that 'African states have engaged in political calculations to avoid the costs and maximise the benefits of cooperation with the ICC.'[92] Moreover, Clark notes the discontent of citizens in the DRC (who echoed the same displeasure as the citizens in Uganda) about state actors escaping criminal liability. One respondent noted, '[t]he most important thing for us is peace…but how can we have peace when we can't trust the government?.'[93] This highlights that if societies are to use international criminal justice as an avenue for peacebuilding, then all perpetrators must be prosecuted and not simply a select few. Unfortunately, the ICTR and ICC have failed to approach their criminal prosecutions collectively, and both tribunals have excluded several individuals who may also bear the greatest responsibility. Therefore, international criminal justice seems to be ill-suited to the African context due to its susceptibility to being manipulated for the benefit of state leaders and to the detriment of local people seeking legal redress and societal cohesion.

'Distant Justice'

A prominent feature of international criminal justice and the courts that administer such law is that they operate outside of the country where the violence has taken place. For example, the ICTR was situated in Arusha in neighbouring Tanzania, whilst the ICC is situated in the Hague in the Netherlands. The reasoning behind international tribunals existing ex-situ is that it will prevent the court from being susceptible to state interference and will thus be able to remain 'free from any political or ideological fetter.'[94] Clark highlights that many African countries that suffer mass atrocities lack the 'infrastructure and personnel to meet the population's legal needs.'[95] In Uganda, '[t]here is almost no judicial presence' in most of the interior, meaning many individuals do not even have access to a justice system at the domestic level.[96] In considering this issue, it may be argued that the creation of the ICC, which has greater financial and administrative resources to carry out prosecutions, could be regarded as a positive

[90] Amnesty International, 'Rwanda, Reports of Killings and Abductions by the Rwandese Patriotic Army' (Amnesty.org 1994) <https://www.amnesty.org/en/documents/afr47/016/1994/en/> accessed 18 February 2022.
[91] Oumar Ba, *States of Justice: The Politics of the International Criminal Court* (Cambridge University Press, 2020), 40.
[92] Sarah Nouwen, *Complementarity in the Line of Fire: The Catalysing Effect of the International Criminal Court in Uganda and Sudan* (Cambridge University Press, 2014).
[93] Phil Clark, *Distant Justice: The Impact of the International Criminal Court on African Politics* (Cambridge University Press, 2018), 120.
[94] Antonio Cassese, 'The Special Tribunal for Lebanon Six Months On: President's Report' in Richard Wilson, *Writing History in International Criminal Trials* (Cambridge University Press, 2011).
[95] Clark (n 51), 113.
[96] ibid.

development in making justice more accessible. However, whilst creating an international criminal justice system that is composed of 'different legal and political cultures of judges and legal staff…'which will'… lend great protection to the goal of delivering justice independently and impartially', it comes at the cost of the court being too far removed from those who most seek its benefit.[97]

Several academic studies have highlighted discontent and ignorance towards the work the ICC conducts from local people on the African continent where atrocities have been committed.[98] Most worryingly, however, are the similarities between the manner in which the ICC is perceived to conduct justice and how justice was conducted during the colonial era.[99] For example, 'when justice happens out of sight, all Congolese worry. The colonials did that…we worry that the ICC won't let us see what's happening.'[100] Furthermore, the six official languages of the ICC are English, French, Spanish, Arabic, Chinese and Russian. Not only does this affirm accusations of the ICC being a neo-colonial power, echoing the assimilation policies employed during the formal colonial period, but the languages make accessibility to justice for many on the African continent increasingly more unattainable. Dominic Ongwen was faced with linguistic challenges when his 1077-page judgement was given to him in a language he could not understand.[101] Whilst it would be unfeasible to accommodate the numerous local dialects present on the African continent, it is problematic to not have even one African dialect among the official languages of the ICC, particularly when nearly all cases heard at the ICC have involved African conflicts. Thus, the inability of many local people to gain access to an international criminal justice system raises significant issues with its compatibility within the African context. Therefore, rather than being a solution to local people in Africa, international criminal justice has only compounded existing issues.

Conclusion

This essay has presented an argument that vehemently upholds the view that international criminal justice has consistently proven to be ill-suited to the African context. The colonial order greatly influenced the hierarchy of judicial practises and how international law developed. Ultimately, it evolved to the detriment of the Global South and calls into question whether international criminal justice could ever have been compatible within the African context. Moreover, the almost exclusive selectivity of the ICC to prosecute crimes committed in Africa, coupled with the power afforded to countries who have not ratified the Rome Statute yet can still influence where investigations are opened, has correctly resulted in the ICC being labelled as a neo-colonial power.[102] The oversimplified definitions of the ideal 'victim' and 'perpetrator' prohibit women to be understood as having agency in armed conflict and also fails to address complex political victims, which condemns them to be solely viewed in their

[97] Gideon Boas, *The Milošević trial: lessons for the conduct of complex international criminal proceedings* (Cambridge, 2007).
[98] Branch (n 42) and Clark (n 51).
[99] Clark (n 51), 130.
[100] Clark (n 51), 131.
[101] Michael Oduor, 'Former LRA leader Dominic Ongwen to appeal ICC ruling, claims mental disability.' (*Africanews*., 2020) <https://www.africanews.com/2021/02/15/former-lra-leader-dominic-ongwen-to-appeal-icc-ruling-claims-mental-disability//> accessed 18 February 2022.
[102] Oumar Ba, 'International Justice and the Postcolonial Condition' [2017] 63 Africa Today 4, 45.

capacity as perpetrators. Furthermore, the lack of recognition and prosecutions pertaining to male victims of sexual violence within international criminal justice means that it is ill-suited to African societies whereby hetero-normative and patriarchal structures are still strongly upheld. Unfortunately, the repeated intervention by the ICC during moments of deep fragility in post-conflict regions severely jeopardised peace processes. Whilst it could be argued that prosecutions may facilitate a more stable form of peace, the continued political instability in regions where indictments have taken place contradicts such claims. Lastly, the locations of and languages used in international criminal justice mechanisms has resulted with victims on the African continent being too far removed from the judicial process to receive any substantial benefit. Most importantly, underpinning all these different aspects is a colonial narrative that has rendered the suitability of international criminal justice within the African context simply unfeasible.

Should Vaccines Be Registered Under Patent Law?

Raneem Alfaleh

Abstract

The recent global challenge of coronavirus enables an overview of how intellectual property (IP) has been used to address global health issues in the vaccine sector. This analysis aims to put into perspective the question of whether patents should be waived during a humanitarian crisis. The article proceeds as follows: the introduction outlines the history of vaccines and patents and provides a definition of vaccines and their contents. Next an overview is given of the recent challenges that were faced rolling out the COVID-19 vaccines. The intellectual property framework is then presented. This research aims to pinpoint whether waiving vaccine patents would be beneficial for the world, especially during a pandemic situation. The research is based on an examination of existing literature. Also, the author relied on a process of critical thinking to provide an overview analysis of the question. Fundamentally, the author reviewed the possible challenges that the pharmaceutical sector might face in relation to vaccination and how waiving a patent might help in delivering vaccines. The research adopted the recent global COVID-19 crisis and the vaccination programme as a case study. The author found that removing vaccine patents during global crises to be logical and proportionate.

Introduction

The physician Edward Jenner introduced the practice of vaccination to the world in a paper published in 1798.[1032] Jenner's observation of the principle behind vaccination and subsequent development of a smallpox vaccine has contributed remarkably to human health worldwide.[1033] Statistics show that vaccines prevent around 2.5 million deaths every year around the globe.[1034] However, despite the vital role played by vaccines in enhancing public health, there are certain challenges facing vaccine roll-out.[1035] For example, there is always a need for new vaccines to help prevent the spread of novel diseases and their variants, but it can take a long time to develop a vaccine.[1036] Moreover, vaccines are expensive to produce, resulting in many potentially effective vaccines being abandoned prematurely.[1037] More importantly, the common issue in the developing world is the problem of gaining access to vaccines, especially those used to combat tropical diseases.[1038]

[1032] 'Learning-History' (*The Jenner Trust*) <https://jennermuseum.com/learning/history> accessed 26 May 2021.
[1033] Hilde Stevens and others, 'Vaccines: Accelerating Innovation and Access' (World Intellectual Property Organisation, 2017) <https://www.wipo.int/edocs/pubdocs/en/wipo_pub_gc_16.pdf> accessed 26 May 2021.
[1034] ibid.
[1035] ibid.
[1036] Petra Oyston and Karen Robinson, 'The current challenges for vaccine development' [2012] 61(7) Journal of Medical Microbiology
<https://www.microbiologyresearch.org/docserver/fulltext/jmm/61/7/889_jmm039180.pdf > accessed 21 May 2021.
[1037] ibid.
[1038] ibid.

In addition, these problems are exacerbated by patents, which grant inventors the right, for a limited period specified by statute, to prohibit anyone from using, producing, advertising, offering for sale or importing their invention without permission.[1039] However, it is argued by some that the power to patent inventions encourages innovation that is motivated by the desire to make money.[1040] Meanwhile, others argue that the waiver of patent rights is sufficient to remove the current obstacles to rapid growth in the supply of vaccines.[1041] The waiving of such rights could be considered especially important during pandemics, when the physical health of populations must take priority over profit.[1042]

In this paper, both the advantages and disadvantages of allowing vaccines to be patented will be considered, to form an opinion on whether vaccines should continue to be protected by patent in the future.

Background

Definition of a 'Vaccine'

Vaccines are biological preparations that consist of damaged or dead versions of a disease-causing organism,[1043] which are introduced into the body to stimulate a natural immune response, thereby enabling resistance to specific diseases.[1044] In other words, a vaccination involves the preparation of dead or inactivated organisms or purified products from which the immune system elicits antibodies to protect against a specific illness (from parasites, viruses, or bacteria).[1045] The World Health Organisation (WHO) defines a vaccine as any preparation designed to create protection against disease through the stimulation of antibody production. Vaccines include, for example, suspensions of dead or attenuated microbes, as well as microorganism products or compounds.[1046] So, vaccines are considered to be important, especially at the present time with different kinds of diseases and epidemics emerging, because vaccines can help in reducing the major impact of pandemics. To protect against the burden of disease, new vaccinations are still needed to battle illnesses such as HIV or malaria, and current vaccines need to be revised and enhanced owing to the introduction of new pathogen strains, such as in the case of influenza.[1047]

[1039] Stevens and others (n 2).
[1040] ibid.
[1041] Scott Burrell, 'Should vaccines be patent protected in a pandemic?' (*Frontier Economics*, 2021) <https://www.frontier-economics.com/uk/en/news-and-articles/articles/article-i8200-should-vaccines-be-patent-protected-in-a-pandemic/> accessed 21 May 2021.
[1042] 'Urgently Waive Intellectual Property Rules for Vaccine' (*Human Rights Watch*, 10 December 2020) <https://www.hrw.org/news/2020/12/10/urgently-waive-intellectual-property-rules-vaccine> accessed 22 May 2021.
[1043] 'Vaccine' (*Cambridge Dictionary*) <https://dictionary.cambridge.org/dictionary/english/vaccine> accessed 22 May 2021.
[1044] Stevens and others (n 2).
[1045] France Innovation Scientifique & Transfert, 'Patent Landscape Report on Vaccines for Selected Infectious Diseases' (World Intellectual Property Organisation, 2012) <https://www.wipo.int/edocs/pubdocs/en/patents/946/wipo_pub_946_3.pdf> accessed 23 May 2021.
[1046] ibid.
[1047] ibid., (n 14).

Content of Vaccines

Water makes up the largest proportion of almost all vaccines,[1048] but the active ingredient of a vaccine is the virus or bacteria against which immunity is sought. This virus or bacteria stimulates the immune system, causing it to produce antibodies to fight the disease.[1049] Inactive ingredients, or 'excipients', are added to the water and active ingredients, and their function is to: either boost the immune response to the vaccine or act as preservatives and stabilisers. These are usually included in very small quantities with some found naturally in our bloodstream. Even so, all excipients are subject to rigorous assessment before they can be included in vaccines, to ensure the substances are safe in the quantities used, with systems in place to monitor their safety on an ongoing basis.[1050]

Thus, each component of the vaccine has a unique purpose, and each ingredient is checked during the production process. All substances are thoroughly evaluated for safety.[1051] A glossary of each of the ingredients is provided below:

> Active Ingredients:
>
> Antigens: Antigens are the core of what makes vaccines work. They are substances that trigger an immune response. In the case of vaccines, the antigen can be the whole inactivated or weakened virus or the bacterium you are trying to trigger a response to; tiny fragments of that pathogen, such as proteins or sugars from it; genetic instructions that tell our own cells how to make those fragments, or weakened viruses used to carry those genetic instructions.[1052]
>
> Inactive Ingredients:
>
> Adjuvants: Vaccine adjuvants boost the immune system's response to the antigen. They can do so by keeping the antigen at the injection site for longer, or by stimulating nearby immune cells. For instance, many vaccines contain aluminium salts, which slow down the release of antigens from the vaccine once it is injected, strengthening, and lengthening the immune response. They also help to stop proteins in the vaccine from sticking to container walls during storage. The amount of aluminium present in vaccines is small, and well below the maximum levels considered safe for humans. Aluminium salts are also commonly added to foods and other medicines.[1053]
>
> Preservatives:
>
> A preservative is occasionally added to vaccination vials, but only if more than one dosage is present. This is done to keep hazardous germs and fungi from contaminating the vaccine every time an individual dosage is withdrawn from it.[1054]

[1048] 'Vaccine Ingredients' (*Vaccine Knowledge Project*, 11 January 2021) <http://vk.ovg.ox.ac.uk/vk/vaccine-ingredients#activeingredients> accessed 23 May 2021.
[1049] ibid.
[1050] Linda Geddes, 'What Ingredients go into a vaccine?' (*Gavi, The Vaccine Alliance*, 21 April 2021) <https://www.gavi.org/vaccineswork/what-ingredients-go-vaccine > accessed 23 May 2021.
[1051] World Health Organisation, 'How are vaccines developed?' (World Health Organisation, 8 December 2020) <https://www.who.int/news-room/feature-stories/detail/how-are-vaccines-developed?topicsurvey=> accessed 23 May 2021.
[1052] Geddes (n 19).
[1053] Geddes (n 19).
[1054] ibid.

Emulsifiers:
Some vaccinations need the addition of emulsifiers such as polysorbate 80, which is often found in food items, to help guarantee that the other components remain suspended in the solution.[1055]

Stabilisers:
Stabilisers, such as sorbitol, which occurs naturally in the body as well as in fruit and berries, are used to preserve vaccine active components from the effects of temperature fluctuations during shipping or storage.[1056]

Residuals:
Vaccines may also contain trace amounts of elements that were employed during the production process but were afterwards removed, for example: cell culture medium, such as egg white; inactivating substances used to destroy viruses or make toxins harmless, such as formaldehyde; or antibiotics used to prevent bacterial infection.[1057]

Consequently, vaccines are one of the most effective medical instruments ever developed, saving more lives than any other medical or public health breakthrough. Prior to COVID-19, it was anticipated that they would prevent 2–3 million lives each year. They do this by carefully simulating a disease with an active substance, generally one that is intrinsically biological, in order to elicit an immune response. However, for these active compounds to be most efficient and safe, they must be combined with other equally important compounds.[1058]

History of Vaccines

Although the science and methods associated with modern vaccination are now highly sophisticated and always evolving in terms of their effectiveness and safety, something akin to vaccination has been practised for centuries. For example, in the sixteenth century, there were accounts of Chinese doctors responding to a smallpox pandemic by grinding up smallpox scabs and administering them to the bodies of healthy people.[1059]

In 1796, a vaccine was developed in a form that resembles what we would recognize today, when an eight-year-old boy was immunised by Edward Jenner using exudate from a cowpox lesion.[1060] In the nineteenth century, the French chemist Louis Pasteur pioneered a vaccination technique by exposing individuals to dead or diluted bacteria, at a time when a better understanding of the process of microbial infection had been developed. Early vaccinations were, however, frequently crudely prepared, and sometimes with serious safety issues. For example, an autoimmune illness was induced in 1 in 3000 vaccinated infants by Pasteur's first rabies vaccine, in which the virus had been cultured using rabbit brain tissue.[1061]

Today, most vaccinations have been in use for decades and are received safely every year by millions of individuals. Every vaccine should be thoroughly and rigorously tested, just

[1055] ibid.
[1056] ibid.
[1057] ibid.
[1058] Geddes (n 19).
[1059] Stevens and others (n 2).
[1060] ibid.
[1061] ibid.

like other medications, to make sure it is safe before it is included in the immunisation programme of a nation.[1062] In the 1980s, advanced molecular biological methods were developed which allowed scientists to enhance the way in which vaccines can imitate diseases. These strategies include the use of viral components, which gave birth to the so-called 'subunit vaccines', as well as the development of recombinant antigens, in which the virus's genes are changed to remove infection while still eliciting immune responses.[1063] The advent of whole genome sequencing and breakthroughs in bioinformatics in the 1990s opened further opportunities for vaccine development. Craig Venter, a biotechnologist, revealed the genome of the first freely living organism in 1995, and towards the end of the twentieth century the genomes of other microorganisms became more easily accessible.[1064] Moreover, researchers also discovered that a variety of additional substances (known as adjuvants) caused a greater immune response when administered in conjunction with an antigen. For a long period, the sole adjuvant in common usage was aluminium salt. More adjuvants have been produced in recent years, each with distinct features targeted to generate a stronger response (i.e., greater effectiveness) and wider immune response to prevent various illnesses. These novel chemicals (e.g., oligonucleotides) have no clinically relevant side effects.[1065] These advances in technology can speed up vaccine research and development (R&D).[1066]

Importance of Vaccines

According to the WHO, the injection of vaccines is known to be the most efficient way of preventing many different diseases in human populations.[1067] At least 20 types of infection, including diphtheria, tetanus, pertussis, pneumonia, and measles, are now prevented by vaccines, saving up to three million lives in human populations every year.[1068] So, a vaccination is a preparation containing dead or inactivated microorganisms (i.e. parasites, viruses, bacteria) or purified compounds generated from them that are intended to stimulate the immune system's response to a specific illness. Prophylactic vaccinations that are used for the treatment of an already existing organism disease in order to protect the system against future infections should be separated from new forms of therapeutic vaccinations (such as those used for various types of cancers). Various forms of vaccinations, such as killed vaccinations, attenuated vaccines, subunit vaccines, conjugate vaccinations, toxoid vaccines, DNA, recombinant vectors, and synthetic vaccines are being developed or are in development.[1069] More importantly, vaccinations can have a favourable influence on future government expenditures by saving lives and significantly decreasing the burden of disease. In many circumstances, the initial expenditure for immunisation might pay for itself multiple times over. Studies have projected how the number of lives saved may affect future government spending, such as in social

[1062] World Health Organisation (n 20).
[1063] Stevens and others (n 2).
[1064] ibid.
[1065] ibid.
[1066] ibid.
[1067] World Health Organisation, 'Vaccines and immunisation: What is vaccination?' (World Health Organisation, 30 December 2020) <https://www.who.int/news-room/q-a-detail/vaccines-and-immunization-what-is-vaccination> accessed 23 May 2021.
[1068] ibid.
[1069] France Innovation Scientifique & Transfert (n 14).

programmes, health care, education, and pensions, as well as in future government tax collections. A research study completed in Egypt projected that the investment expense of a rotavirus vaccination programme for children would be completely offset by the time they reached the age of 22.[1070]

Overview of Recent Challenges to Vaccine Roll-out

Vaccines have enhanced public health around the world. This is because, by being vaccinated, people protect themselves from contracting or spreading the diseases against which they are immunised.[1071] However, there are still significant hurdles ahead. For example, by 2015, 126 nations had achieved at least 90% coverage of the third dose of the diphtheria-tetanus-pertussis (DTP3) vaccination.[1072] Unfortunately this meant that around 19.4 million children were still at risk.[1073] Therefore, the challenge that many people face worldwide consists of the real threat of deadly disease when vaccines are not accessible to them.[1074] This is especially the case in the world's poorest countries, where governments and individuals cannot always afford vaccines: according to a 2017 report, 'The cost of a specific vaccination package has risen by 2,700 percent over 10 years: from USD 1.37 in 2001 to USD 38 in 2011.'[1075] Moreover, newer vaccinations, in particular, such as rotavirus and pneumococcal conjugate vaccines which target the most prevalent causes of sickness and mortality in children – diarrhoea and respiratory infection – are not widely available.[1076] To illustrate this point, with the current COVID-19 pandemic the world has been faced with huge challenges. At the present time, it has been reported that 'Sixteen countries have extended the timing between vaccine doses to provide the first dose to as many people in the priority groups as possible. The timing between the first and second dose varies by country and by vaccine product.'[1077] Moreover, according to UNICEF data, only 43% of COVID-19 vaccine manufacturing capacity is presently used for authorised vaccinations. This is due to pharmaceutical companies' desire to safeguard their intellectual property – the patented formulas and the technology used to create the vaccine – and not share with governments and manufacturers the capability to create additional doses.[1078] Also, according to a report from the European Centre for Disease Prevention and Control, the majority of nations (15 of 23 reporting countries, or 65 percent) indicated that the biggest barrier to vaccination deployment is a lack of vaccine supplies.[1079] Not only that, with the COVID-19 vaccine there was significant hesitation in taking the vaccine among people, with

[1070] Stevens and others (n 2).
[1071] ibid.
[1072] ibid.
[1073] ibid.
[1074] ibid.
[1075] ibid.
[1076] ibid.
[1077] European Centre for Disease Prevention and Control, 'Overview of the implementation of COVID-19 vaccination strategies and deployment plans in the EU/EEA' (European Centre for Disease Prevention and Control, 14 June 2021) <https://www.ecdc.europa.eu/en/publications-data/overview-implementation-covid-19-vaccination-strategies-and-deployment-plans> accessed 23 May 2021.
[1078] Will Meyer, 'If we don't waive vaccine patents, thousands of people will needlessly die' (*Business Insider*, 4 May 2021) <https://www.businessinsider.com/vaccine-patents-ips-monopoly-vaccination-process-us-biden-africa-2021-2?r=US&IR=T> accessed 23 May 2021.
[1079] European Centre for Disease Prevention and Control (n 46).

eleven nations observing a rise in vaccination hesitancy, primarily due to a lack of acceptance with regard to the Vaxzevria vaccine.[1080] Furthermore, with the recent challenges with the COVID-19 vaccination, vaccine certificates for COVID-19 will be issued by the majority of responding nations, and several have begun conversations about the future use of these certificates, for example relating to travel, tourism, facilitating the easing of non-pharma interventions and enabling admission to certain locations or events.[1081] However, there are some distribution issues that still need to be addressed and key actions might help to solve the shortages in the delivery of vaccinations. Therefore, for this scenario, the WHO has developed a framework that outlines four cumulative criteria for ensuring the long-term supply of medicines: '1. rational selection 2. affordable prices 3. sustainable financing, and 4. reliable health and supply systems.'[1082]

Vaccines and the Intellectual Property (IP) Framework

Introduction

Vaccines and vaccine-related technology are subject to a variety of intellectual property rights, including patents.[1083] Moreover, vaccines, as with medicines and other technologies, are protected by patents, which offer legal protection against copying.[1084] Patents provide producers the right and means to make more money with their discoveries – an incentive to stimulate innovation.[1085] However, according to the WHO, the vaccine itself is not protected under the IP and patent rights because it contains multiple factors, as has been shown previously, such as antigens, adjuvants and excipients (see Section 2.2 Content of Vaccines above). Therefore, each one of those factors is instead protected under the IP and has its own patent right.[1086] In other words, vaccines contain multiple levels and each one is protected as an individual factor and has a patent right so each of these levels are protected and, at the end, they combine to produce the vaccines which are themselves not protected.[1087] Thus, patents can cover the formulation of the vaccination as well as the combination of therapeutic components. Patents might also be found on the vaccine administration equipment, such as in relation to an injectable delivery system or a capsule designed to release the substance in a specific location of the human body. Third parties may be prevented from exploiting clinical trial data submitted to regulatory bodies to get marketing authorization to produce competitive products, if the clinical test data is protected.[1088]

[1080] ibid.
[1081] ibid.
[1082] World Health Organisation, 'Essential medicines and health products' (World Health Organisation, 21 December 2016) <https://www.who.int/medicines/areas/access/en/> accessed 23 May 2021; Stevens and others (n 2).
[1083] Stevens and others (n 2).
[1084] Dharshini David, 'Covid: The Vaccine Patent Row Explained', (*BBC*, 6 May 2021) <https://www.bbc.co.uk/news/business-57016260> accessed 12 July 2021.
[1085] ibid.
[1086] Martin Friede, 'Intellectual Property and Licence management with respect to Vaccines' (World Health Organisation, 2010) <https://www.who.int/phi/news/Presentation15.pdf?ua=1> accessed 24 May 2021.
[1087] ibid.
[1088] Stevens and others (n 2).

On the other hand, according to the WHO, for some basic vaccines which were produced at least 20 years ago, such as 'D, T, Pw, Pa, HepB, HiB, IPV, OPV, measles, mumps, rubella, yellow fever', it is 'impossible' that intellectual property can constitute a barrier to their manufacturing, use, or sale as these are considered 'classical' formulations.[1089] However, there are some exceptions highlighted by the WHO, such as 'Improved formulations – Combinations, adjuvants, doses, delivery routes' and 'Improved processes for manufacture.'[1090]

What is a Patent?

Intellectual property (IP) captures the ideas, literary and artistic works, compositions, marks, titles and pictures used in trade.[1091] 'Patents, copyright, trademarks and trade secrets',[1092] for example, are legal mechanisms that allow individuals to be credited for or to obtain financial advantage from their inventions or ideas.[1093] Trademarks, which assist to create a relationship between products or services and a certain organisation or individual creating or delivering them, may be used to protect a vaccine's brand name. Copyright protects the expression of ideas and extends to vaccination explanatory materials and designs. Finally, trade secrets safeguard knowledge that inventors and corporations prefer not to publish, knowledge that is non-codifiable, or knowledge that does not fulfil patentability standards.[1094] Therefore, IP rights, as with patents, can also be used to gain control over vaccine manufacturing and distribution, for example, through licensing. This control can assist in ensuring the quality and safety of vaccines. Quality control is an important component in shaping public opinion of a vaccine's quality and effectiveness.[1095] Similarly, trademarks serve as the foundation of quality assurance systems, allowing an inventor to capitalise on patients' faith in the protected vaccination. In this context, it is crucial to highlight that public acceptance of a vaccine can be critical for immunisation programme efficacy.[1096] Thus, the debate surrounding intellectual property rights and vaccines should not begin and end with the application of one IP right to a vaccine.[1097] The debate should include considerations of different IP rights relevant to a vaccine and how they may be used in an integrated way within a plan targeted at helping the vaccine's development and dissemination. Such an approach to IP rights for vaccines allows the integrated rights to be considered in light of the justifications for protecting vaccines with IP rights, as well as issues relating to specific IP rights for vaccines, such as compulsory licence regimes, available humanitarian purpose IP credits, and so on.[1098]

The IP framework seeks to promote an atmosphere in which imagination and invention can thrive by finding the right balance between the needs of innovators and the wider public

[1089] Friede (n 55).
[1090] ibid.
[1091] World Intellectual Property Organisation, 'What is Intellectual Property?' (World Intellectual Property Organisation) <https://www.wipo.int/about-ip/en/> accessed 23 May 2021.
[1092] ibid.
[1093] ibid.
[1094] ibid.
[1095] Stevens and others (n 2).
[1096] ibid.
[1097] Karen Durell, 'Vaccines and IP Rights: A Multifaceted Relationship' in Sunil Thomas (ed) *Vaccine Design: Methods in Molecular Biology* (vol 1404 Humana 2016). <https://doi.org/10.1007/978-1-4939-3389-1_52> accessed 23 June 2021.
[1098] ibid.

interest.[1099] Vaccine inventions enjoy a range of IP protection, including patents, trademarks, copyrights, and trade secrets.[1100] In particular, a patent is an IP right that can be sought in relation to vaccine technologies.[1101] Patents protect the composition of a vaccine, including the mix of pharmaceutical ingredients.[1102] So, patents can cover the formulation of the vaccination as well as the combination of therapeutic components. Patents may also be present on the equipment used to administer vaccines including, for example, injectable delivery systems or capsule systems designed to release the product in the human body rather than through injection. Third parties may also be prohibited from using clinical trial data once it has been submitted to regulatory bodies for marketing authorization; if the clinical test data is protected, it cannot be used by others to manufacture competitive products.[1103] So, in the case of vaccines, one of the major functions of intellectual property rights is to promote the large R&D investments necessary for their creation. Vaccine innovation may not emerge in the absence of this incentive. Furthermore, intellectual property rights such as patents can allow control over vaccine manufacturing and distribution, for example, through licensing. This control can assist to assure the quality and safety of vaccines.[1104] Therefore, control of quality is a crucial aspect for public perception of vaccination quality and efficacy. Similarly, trademarks serve as the foundation of quality assurance systems, allowing an inventor to capitalise on patients' faith in the protected vaccination. In this context, it is crucial to highlight that public acceptance of a vaccine can be critical for immunisation programme efficacy.[1105]

Overview of Patent Registration in Saudi Arabia and the United Kingdom

As mentioned above, patents form a part of IP protection frameworks. Therefore, the author will briefly present a comparison between two jurisdictions in patent registration. In the Kingdom of Saudi Arabia, for example, a patent may be registered by submitting a patent file to the Saudi Patent Office[1106] or a patent application may be filed with the Patent Office of the Gulf Cooperation Council (GCC).[1107] The protection for forms of patents in Saudi Arabia is 20 years from the filing date.[1108] However, there are some terms where a patent might not be granted under the Saudi law. For example, Article 4 of the Saudi Arabian Patent Law states the following:

> (a) The protection document shall not be granted if its commercial exploitation violates the Sharia (Islamic) Law.

[1099] World Intellectual Property Organisation (n 60).
[1100] Stevens and others (n 2).
[1101] World Intellectual Property Organisation (n 60).
[1102] ibid.
[1103] ibid.
[1104] ibid.
[1105] ibid.
[1106] European Patent Office, 'FAQ – Saudi Arabia (SA)' (European Patent Office) <https://www.epo.org/searching-for-patents/helpful-resources/asian/faq.html> accessed 23 May 2021.
[1107] ibid.
[1108] ibid.

> (b) The protection document shall not be granted if its commercial exploitation is harmful to life, to human, animal, or plant health, or is substantially harmful to the environment.[1109]

However, there are some inventions which cannot be patented, as stated in Article 45 of the Saudi law as follows:

> In the application of provisions of this Law, the following shall not be regarded as inventions:
>
> (a) Discoveries, scientific theories and mathematical methods.
> (b) Schemes, rules and methods of conducting commercial activities, exercising pure mental activities or playing a game.
> (c) Plants, animals and processes - which are mostly biological - used for the production of plants or animals, with the exception of microorganisms, non-biological and microbiology processes.
> (d) Methods of surgical or therapeutic treatment of human or animal body and methods of diagnosis applied to human or animal bodies, except for products used in any of these methods.
>
> The exclusion also applies to computer programs and any other copyright work.[1110]

In contrast, a UK patent must be registered by submitting a patent application to the UK's Intellectual Property Office (IPO).[1111] The UK law has specified that for an invention to be patented, it must be all of the following: 'something that can be made or used; new; inventive - not just a simple modification to something that already exists.'[1112] On the other hand there are some types of invention which cannot be patented, and these include:

1) literary, dramatic, musical or artistic works;
2) a way of doing business, playing a game or thinking;
3) a method of medical treatment or diagnosis;
4) a discovery, scientific theory or mathematical method;
5) the way information is presented;
6) some computer programs or mobile apps, and;
7) 'essentially biological' processes like crossing-breeding plants, and plant or animal varieties.[1113]

Moving on to the validation period of a patent, in Saudi Arabia a patent is valid for fifteen years from the decision of grant and can be extended for another five years. A patent is also subject to annuity payments, which are to be paid only one year after the decision of the grant of the patent.[1114]

[1109] ibid.
[1110] ibid.
[1111] GOV.UK, 'Patenting your invention' (GOV.UK) <https://www.gov.uk/patent-your-invention/apply-for-a-patent> accessed 26 May 2021.
[1112] ibid.
[1113] ibid.
[1114] Abu-Ghazaleh Intellectual Property, 'Saudi Arabia: Summary of the Patent Registration System in Saudi Arabia' (*Mondaq*, 23 December 1999) <https://www.mondaq.com/saudiarabia/trademark/4026/summary-of-the-patent-registration-system-in-saudi-arabia> accessed 23 May 2021.

Whereas in the United Kingdom, once a patent is granted by the IPO, and the application has been published and a certificate delivered, it is the customer's duty to determine how long the patent lasts; if the patent is renewed every year it may last up to 20 years from the date of filing the application. The benefit is that renewals commence only on the fourth anniversary of the filing date, and are only subject to a charge of £70, which rises every year to £610 in year 20, therefore meaning there is not a big amount to pay in the beginning.[1115]

Patenting of Vaccines during the COVID-19 Crisis

As the COVID-19 epidemic spreads throughout the world, thousands of patients urgently require access to inexpensive medications. Based on previous experience with therapies for other life-threatening diseases, there is concern that access to any future vaccinations and therapies may be hampered by patents, leading to unaffordable high pricing.[1116] Therefore, there is a debate that, in some circumstances, the patent on vaccines should be waived.[1117] Given the uneven situation, the poor nations have proposed that intellectual property rights in relation to COVID-19 treatments be suspended. They believe that patenting vaccines and other treatments concentrates the supply in the hands of wealthy countries, excluding poorer countries that have struggled to acquire access to them in the past.[1118] According to Steve Bates, chief executive of the Bioindustry Association, a trade group for innovative life sciences based in the United Kingdom, eliminating intellectual property rights is 'not a panacea.'[1119] In his estimation, handing a recipe book to a country's government without the ingredients, protections, infrastructure, and large staff with the high skills required to administer safe and efficient vaccinations would not quickly offer aid to all the people who need it.[1120]

The World Trade Organisation (WTO) has not yet granted the waiver, but if it does, analysts estimate that the world will not see increased capacity until at least 2022.[1121] For example, during a pandemic, such as the current COVID-19 pandemic, it could be determined that urgent action is required from pharmaceutical corporations to share IP with other countries.[1122] On the other hand, the International Federation of Pharmaceutical Manufacturers & Associations (IFPMA) said in a statement that 'Waiving patents of COVID-19 vaccines will not increase production nor provide practical solutions needed to battle this global health

[1115] 'UK Patents: Five things you should know' (*MPA*) <https://mpa.co.uk/what-we-do/tax-services/patent-box/uk-patents-five-things-you-should-know/> accessed 24 May 2021.
[1116] Olga Gurgula, 'Strategic Patenting by Pharmaceutical Companies – Should Competition Law Intervene?' [2020] IIC – International Review of Intellectual Property and Competition Law 51 <https://link.springer.com/article/10.1007/s40319-020-00985-0.> accessed 23 May 2021.
[1117] Michele Boldrin, David Levine and Flavio Toxvaerd, 'Should patents on Covid-19 vaccines be waived?' (*Economics Observatory*, 14 May 2021) <https://www.economicsobservatory.com/should-patents-on-covid-19-vaccines-be-waived> accessed 23 May 2021.
[1118] Burrell (n 10).
[1119] Gareth Iacaobucci, 'Covid-19: How will a waiver on vaccine patents affect global supply?' (10 May 2021) 373 The BMJ: British Medical Journal (Online) <https://www.bmj.com/content/373/bmj.n1182> accessed 23 May 2021.
[1120] ibid.
[1121] ibid.
[1122] James Paton and Tim Loh, 'Countries Need Vaccines Now, and Patent Waivers Won't Deliver Them' (*Bloomberg,* 7 May 2021) <https://www.bloomberg.com/news/articles/2021-05-07/countries-need-shots-now-and-patent-waivers-won-t-deliver-them> accessed 23 May 2021.

crisis.'[1123] However, it has been claimed that a waiver of IP rights is allegedly permitted for more readily shareable vaccination technologies. This would mean that in nations who have the production capability, generic or otherwise non-licensed manufacturers might commence manufacturing (such as in India and Brazil).[1124] Whereas, on the other side, the European Union president has stated that Europe has committed billions to help to produce the first COVID-19 vaccines in the world and to provide a worldwide shared benefit.[1125] However, if there were no strings or conditions attached to the original investment when it was converted into patented ideas, there is no 'global common good' when such vaccine commodities can be appropriated and monopolised by patents.[1126]

However, several firms appear to already own significant patents covering some of the vaccines.[1127] In a certain case (Moderna), the manufacturer has submitted a follow-up application to a prior patent application that included their vaccination technique for additional coronaviruses. In other circumstances (Novovax), the company's vaccine is based on private technology that was already patented. In another situation (AstraZeneca), the corporation appears to have adopted both approaches, recently filing a follow-on patent application from a prior patent that included the technology utilised to build its current COVID-19 vaccine. In another situation (J&J), the corporation may end up depending on patents protecting vaccine production, or more specifically, big quantities of vaccine.[1128] Therefore, most of the companies will be using either an old patent right or a newer one in order to protect their product.[1129] Meanwhile Bill Gates, whose private foundation has given to and partnered with Covax (a vaccine distribution mechanism in the developing world), has openly stated that vaccine manufacturers should not relinquish their patents.[1130] Yet, the main objection from vaccine manufacturers and their home nations is that just surrendering patents would not be of practical benefit. They argue it would be like sending someone a recipe without the ingredients or directions, as the patent protects the bare bones of the blueprint but not the exact manufacturing method. That is critical in this case, since mRNA vaccines, such as those made by Pfizer and Moderna, are a novel breed, and only a few individuals know how to manufacture them.[1131]

[1123] Michael Safi, 'Covid vaccines: what is patent waiving and will it solve the global shortage?' *The Guardian* (London, 6 May 2021) <https://www.theguardian.com/world/2021/may/06/covid-vaccines-what-is-patent-waiving-and-will-it-solve-the-global-shortage> accessed 23 May 2021.
[1124] Burrell (n 10).
[1125] Siva Thambisetty, 'Vaccines and patents: how self-interest and artificial scarcity weaken human solidarity' (*London School of Economics*, 9 February 2021) <https://blogs.lse.ac.uk/politicsandpolicy/vaccines-and-patents/> accessed 24 May 2021.
[1126] ibid.
[1127] Zachary Silbersher, 'Which patents cover the COVID-19 vaccine candidates for Moderna, AstraZeneca, J&J and Novovax?' (*Markman Advisors,* 21 July 2020) <https://www.markmanadvisors.com/blog/2020/7/21/which patents-cover-the-covid-19-vaccine-candidates-for-moderna-astrazeneca-jampj-and-novovax> accessed 23 May 2021.
[1128] ibid.
[1129] ibid.
[1130] Meyer (n 47).
[1131] David (n 53).

An Overview of Patents in the Pharmaceutical Sector

The pharmaceutical sector is one of the most sophisticated in the world. It is distinguished by strict governmental control and, at times, by the clash of interests between the pharmaceutical industry and society. It also involves a variety of parties, including originators, marketing authorization Organisations, generic businesses, doctors, pharmacies, and patients. Each of these parties contributes to the lengthy and complex process of converting a chemical substance into an effective and economical drug that is then prescribed, distributed, and consumed.[1132] Two essential actors – generic firms and originating firms – play critical roles in these intricate connections. Generic firms help society by providing cheaper counterparts of the originators' medications, resulting in lower medication prices and more access to inexpensive medications.[1133] When the interests of these two actors are balanced, the advantages to society are maximised, as it obtains novel and better treatments, as well as timely access to generic pharmaceuticals. However, if the balance shifts in favour of one of the players, society suffers as a result of a lack of access to either new or economical medications. As a result, both pharmaceutical innovation and generic competition must be appropriately rewarded and safeguarded.[1134]

Therefore, drug companies would prefer to be allowed to patent medicines, including vaccines, because the money generated helps to fund more research as costs of trials are considered to be very expensive.[1135] There are various reasons for the high level of expense for pharmaceutical companies, such as the intricacy of drug research and development, as well as the costly and time-consuming regulatory procedures required.[1136] Most importantly, pharmaceutical firms rely substantially on intellectual property rights, particularly patents, to secure their enormous efforts and investments.[1137] According to Drugwatch, 'new medicines that treat illnesses affecting the respiratory system are the most expensive trials to run and average about $115.3 million.'[1138] Therefore, as a result of the high cost to develop new drugs, 'In 2012, pharmaceutical companies paid $39 billion for trials, while NIH [the US National Institutes of Health] paid $31 billion.'[1139] Therefore, the pharmaceutical sector relies heavily on patent protection. As medication research is both costly and time-consuming, pharmaceutical firms want enough protection to recoup their investments and produce revenues, allowing them to fund more R&D.[1140] Also, industry groups are concerned that without access to all of the know-how and parts, a waiver might result in quality, safety, and effectiveness concerns, as well as counterfeits. These groups point out that Moderna has previously said that it will not pursue anybody discovered to be infringing on their patent – though such a scenario has yet to occur.[1141]

[1132] Gurgula (n 85).
[1133] ibid.
[1134] ibid.
[1135] Michelle Llamas, 'Big Pharma's role in clinical trials' (*Drugwatch*, 24 April 2015) <https://www.drugwatch.com/featured/clinical-trials-and-hidden-data/> accessed 23 May 2021.
[1136] Gurgula (n 85).
[1137] Gurgula (n 85).
[1138] ibid.
[1139] Llamas (n 104).
[1140] Gurgula (n 85).
[1141] David (n 53).

There is another view that waiving patents would not solve the issue of the shortage supply in vaccines, especially during the COVID-19 crisis, as it seems doubtful that lifting patent protection would be sufficient to overcome the present hurdles that are limiting quicker increases in supply.[1142] As it is, every vaccine producer has a strong incentive to maximise the amount of vaccine accessible. There is substantial surplus demand around the world and maintaining an abundant supply of dosages in the short term is likely to result in major reputational advantages for the firms involved. Developing relationships with potential clients now increases the likelihood of winning more profitable contracts in the future.[1143] Despite massive worldwide demand — and strong incentives to expand output – severe supply constraints exist. This implies that the hurdles to growing supply go beyond patent ownership and intellectual property rights in the relevant technology.[1144]

Indeed, production capabilities and regulatory approvals are major obstacles for every business to overcome.[1145] For instance, initial approval from the European Medical Association (EMA) for AstraZeneca to manufacture vaccines was granted in January 2021 – this covered manufacturing at sites in the UK, Belgium and the United States.[1146] However, a fourth site has been approved for manufacturing in the Netherlands – approval was granted in March 2021.[1147] Regulatory authorities in several countries have halted the deployment of some vaccinations due to health concerns. Similar difficulties would arise in underdeveloped countries, as well as for any firm wishing to supply such areas.[1148] Because this may undermine public trust in the safety and effectiveness, and as a result, have a negative effect on the uptake rates of vaccinations, many would not regard the expansion of production at the expense of safety as an acceptable trade-off.[1149]

An Alternative Perspective: Should Vaccine Patents be Waived?

The issue of exorbitant medicine pricing is not new. Long before the pandemic, increasing healthcare expenses had posed a severe threat to the affordability and accessibility of medications for society.[1150] The success of an enterprise depends on its business performance in competitive marketplaces. Therefore, companies must innovate in order to compete in performance via providing higher quality and a broader selection of new and enhanced goods and services. Given the significance of safeguarding innovation as the major driver of economic growth, nations have put in place several measures to guarantee an appropriate environment for the development of innovation. They include providing property rights in the form of patents to the outcomes of invention and the use of competition law to boost dynamic

[1142] Burrell (n 10).
[1143] ibid.
[1144] ibid.
[1145] ibid.
[1146] ibid.
[1147] European Medicines Agency, 'Increase in vaccine manufacturing capacity and supply for COVID-19 vaccines from AstraZeneca, BioNTech/Pfizer and Moderna' (European Medicines Agency, 26 March 2021) <https://www.ema.europa.eu/en/news/increase-vaccine-manufacturing-capacity-supply-covid-19-vaccines-astrazeneca-biontechpfizer-moderna> accessed 24 July 2021.
[1148] ibid.
[1149] ibid.
[1150] Gurgula (n 85).

competition.[1151] The promotion of invention, which is the engine for economic growth and expansion, is one of the key reasons for the patent system. The patent system follows this purpose by providing patent proprietors with a time limit for exclusive rights in order to reward their original work and to promote subsequent invention. Therefore, IP rules, and in particular patents, are considered as a crucial component of undistorted domestic competition.[1152] To illustrate, patents and competition laws complement each other, since on the one hand current competition places demands on businesses, driving them to innovate, the so-called 'stick,' while on the other hand, patent law offers 'the carrot.' These two legislative entities are viewed as supplementary measures towards promoting an effective marketplace and dynamic innovation competitiveness in the long term. Their function is to encourage innovation and ensure the competitive use of it.[1153] As the European Commission has highlighted, 'intellectual property rights and competition are required.'[1154] Moreover, the EU competition law prohibits methods which decrease innovation incentives for both 'pioneers' and subsequent innovators. This is codified in Article 102 TFEU (Treaty on the Functioning of the European Union), which bans abuses such as, among other things, restricting technical advancement.[1155]

In *AstraZeneca*, for example, the European General Court determined that the company's practice of abusing the patent system had the potential to reduce its incentives to develop and was therefore anticompetitive.[1156] Courts ruled that the proprietors of IP rights misused their dominant positions in the cases of *Magill* and *Microsoft* by impeding their prospective competitors' innovation.[1157] To illustrate, 'Following the rationale of the European General Court's statement in AstraZeneca, the practice of the originator that extends its market monopoly by relying on the patent system 'potentially reduces the incentive to engage in innovation, since it enables the company in a dominant position to maintain its exclusivity beyond the period envisaged by the legislator.' Such practises, according to the Court, act 'contrary to the public interest.'[1158] However, some pharmaceutical corporations may utilise the patent system for a different goal, such as obtaining several secondary patents that construct multi-layer protection around their profitable drugs to hinder merit-based competition. Such selective patenting enables pharmaceutical corporations to strengthen their monopolies and hence continue to charge exorbitant medicine costs.[1159] Such a situation can arise because a patent grants its owner the legal right to prevent others from utilising their creation for a set number of years in return for making the innovation publicly available.[1160]

Therefore, patents can limit the ability of new manufacturers from less privileged nations to enter vaccination markets at any stage of the regulatory process, from preclinical research and development to scale-up, formulation, and licensure in the markets of choice. According to a recent research study, several firms in developing nations complain that the

[1151] ibid.
[1152] ibid.
[1153] Gurgula (n 85).
[1154] ibid.
[1155] ibid.
[1156] ibid.
[1157] ibid.
[1158] ibid.
[1159] ibid.
[1160] ibid.

requirement to handle patents raises transaction costs and lengthens development timeframes.[1161] An exception that would suspend drug patents for vaccinations would make it possible for underdeveloped countries to enter the market. In the case of vaccinations, more than simply a waiver will be required; technical know-how and further knowledge transfer may also be required.[1162] Therefore, in terms of any regulatory processes ranging from preclinical R&D through scaling, formulation and licensing in the market of choice, patents might affect the capacity for new producers from poor nations to participate in vaccination markets. According to a recent research study, several manufacturers operating in developing nations say that the requirement to handle patents raises transaction costs and extends development timetables.[1163] Moreover, patents on vaccines and key technology for vaccine development are just one of several variables influencing vaccine innovation and availability in underdeveloped nations. According to a recent study conducted in Brazil, the most significant barriers to boosting vaccination coverage include insufficient regulatory frameworks and processes, as well as low levels of investment in local capacity, human resources, technology, and logistics.[1164] The case for systemic adjustments in incentives for health technology via intellectual property legislation is not new. Therefore, the COVID-19 pandemic has re-energized the long-running debate over worldwide access to health technology by calling to light the problematic connection between intellectual property law and global health.[1165]

In their report on access to medicines in 2016, the United Nations Secretary-General recognized disruptions in the legal aspects, policy and practice of the right to health and international trade, especially in terms of conflicts between the rules of intellectual property and the aims of public health.[1166] Thus, during the current pandemic, almost 51% of the global COVID-19 vaccines – which accounts for around 3.7 billion doses – has been taken by the richer and developed countries.[1167] It should be noted that these countries constitute only 14% of the global population.[1168] This has resulted in a major shortage of accessible dosages for purchase. As a result, the number of doses obtained by poorer nations – who rely heavily on WHO non-profit programmes to purchase vaccinations on their behalf – is insufficient and only able to offer a full course of therapy to a tiny proportion of their people.[1169] Despite the fact that governments largely contributed to the research and development of the top vaccines, companies may still benefit from medications that governments sponsor due to intellectual property regulations.

In the United States, this is owing to the Bayh-Dole Act of 1980, which allows firms to patent government-funded drugs. According to the economist Dean Baker, 'the amount of

[1161] Stevens and others (n 2).
[1162] Thambisetty (n 94).
[1163] Stevens and others (n 2).
[1164] ibid.
[1165] Siva Thambisetty and others, 'The TRIPS Intellectual Property Waiver Proposal: Creating the Right Incentives in Patent Law and Politics to end the COVID-19 Pandemic' (24 May 2021) LSE Legal Studies Working Paper 06/2021 <https://ssrn.com/abstract=3851737> accessed 13 July 2021.
[1166] ibid.
[1167] Burrell (n 10).
[1168] Anthony D So and Joshua Woo, 'Reserving coronavirus disease 2019 vaccines for global access: cross sectional analysis' (15 December 2020) 371 The BMJ: British Medical Journal (Online) <https://www.bmj.com/content/371/bmj.m4750> accessed 21 January 2022.
[1169] Burrell (n 10).

money shifted from the rest of us to those in a position to gain from IP amounts to more than $1 trillion yearly.'[1170] Moreover, the current lack of COVID-19 vaccines, medical equipment, medicines, and diagnostics (referred to under the generic term 'health technologies') is due to the structure of global intellectual property law as embodied in the WTO's Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS). The TRIPS Agreement mandates Member States, subject to standard assessments on novelty, creativity and industrial applicability, to make patents accessible for any product or process of discovery in all fields of technology, without discrimination.[1171] Patents and patent rights must also be available without prejudice as to the site of invention and whether items are imported or locally manufactured.[1172] Although equitable access to vaccines is in the moral, political, and economic interests of the global community and necessitates global solidarity, the phenomenon of COVID-19 'vaccine nationalism' has highlighted the misalignment of current legal and financial incentives to produce and distribute vaccines equitably.[1173] The crisis also demonstrates the failure of high-income countries (HICs) to keep their promise made during TRIPS negotiations in 1994, that by agreeing to the TRIPS terms, lower and middle income countries (LMICs) would benefit from technology transfer and the development of productive capacity. The present issue reveals limitations not only in the way global catastrophes might be addressed, but also shortcomings within the international patent agreement itself.[1174]

Therefore, the epidemic has worsened global inequities that already existed, none more so than in vaccine manufacturing and distribution.[1175] As of 14 May 2021, the United States and the United Kingdom had immunised around half of their adult populations against COVID-19 with at least one vaccination dosage. Israel had already surpassed this target, having vaccinated 63 percent of its people with a single shot. The EU nations have been catching up: by this date, the cumulative uptake of one vaccination dose for adults over the age of 18 was 36% among EU/EEA countries. However, billions of people in Asia, Africa, and Latin America remain unvaccinated, and many of them have little possibility of obtaining a COVID-19 vaccine in 2021, 2022, or even 2023.[1176] Therefore, some countries have supported the idea of waiving patents in order to protect their national public health. For example, in the United States the Biden administration has indicated that it will accept the elimination of patents on COVID vaccinations but not on therapies or other disease-fighting technology. If the WHO approves the waiver, businesses creating COVID-19 vaccines across the world will be able to do so without fear of being sued by another Organisation that already has the patent on the product.[1177] India and South Africa called for a waiver that would 'continue until widespread vaccination is in place globally, and the majority of the world's population has developed immunity.'[1178] This call has not been heeded at successive WTO meetings, though it received

[1170] Meyer (n 47).
[1171] World Trade Organisation, 'Overview: the TRIPS Agreement' (World Trade Organisation) <https://www.wto.org/english/tratop_e/trips_e/intel2_e.htm> accessed 12 July 2021.
[1172] ibid.
[1173] Thambisetty (n 134).
[1174] ibid.
[1175] ibid.
[1176] ibid.
[1177] Safi (n 92).
[1178] Thambisetty (n 134).

a recent boost in May 2021 via US support for a narrower IP waiver applying only to vaccines against COVID-19. Although sometimes referred to in shorthand as a 'patent waiver', in both its original and revised forms the India/South Africa proposal is in fact a broad package, applicable to diagnostics, treatments and vaccines. It is currently co-sponsored by 62 WTO countries (including India and South Africa). The waiver would apply 'in relation to prevention, containment or treatment of COVID-19', covering not only the temporary waiver of patents (and, where relevant, copyrights) internationally, but also, crucially, the sharing of IP under the umbrella of 'undisclosed information' such as trade secrets and 'know-how.'[1179] In principle, this kind of 'sharing' is not new. The 2011 WHO Pandemic Influenza Preparedness (PIP) Framework makes explicit reference to technology transfer, albeit in the somewhat limited context of benefit sharing (in return for receiving biological materials), and it offers language that is short of a legal obligation.[1180] However, some of the text bears repeating here. Section 6.13.4 states as follows:

> *'Influenza vaccine manufacturers who receive PIP biological materials may grant, subject to any existing licensing restrictions, on mutually agreed terms, a nonexclusive, royalty-free licence to any influenza vaccine manufacturer from a developing country, to use its intellectual property and other protected substances, products, technology, know-how, information and knowledge used in the process of influenza vaccine development and production, in particular for pre-pandemic and pandemic vaccines for use in agreed developing countries.'*[1181]

In general, IP's legal incentives are organised in such a way that ideas that can be easily replicated or reverse-engineered tend to be patented. If such an invention is lacking in patent protection, then it could be quickly reverse-engineered or copied by a rival in the market. On the other hand, if an innovation is truly difficult to reproduce, it may make more strategic business sense to keep that innovative information as a trade secret — and perhaps then get longer protection than a patent permits.[1182] Companies who create and promote such a product rely on the fact that no one can easily 'read' it.[1183] Furthermore, many entrepreneurs, health policy specialists and government officials worry that the coronavirus will continue to spread and mutate if the vaccination campaign drags out because the vaccine capacity is dictated by intellectual property regulations rather than by necessity.[1184] Continuing mutational change and propagation would have a harder influence on the international economy than the reduction of these artificial access barriers that protect the rich (and the global institutions involved) and would enable faster vaccination.[1185] The consequences of such short-sighted financial-led thinking are being felt in real time, as new outbreaks of the virus ravage nations such as Brazil

[1179] ibid.
[1180] ibid.
[1181] ibid.
[1182] ibid.
[1183] ibid.
[1184] Meyer (n 47).
[1185] ibid.

and India.[1186] It is believed that the TRIPS waiver is a crucial legal instrument to enable a radical increase in the capacity of COVID-19 vaccines.

Therefore, a temporary waiver of all applicable IP will be provided under the TRIPS waiver as anticipated by the India/South Africa proposal. This includes, but is not limited to, patents. The advantage of the universal exemption of patents on COVID-19 vaccines and health technologies is that, given the above problems of disclosure, transparency and overlapping of patents, producers could be free to operate without litigation and fear that exported vaccines could be seized and accused of alleged patent infringement.[1187] Thus, with the scope of the global public health problem and the concept of the patent thicket — a dense maze of trade secrets hidden behind an information interface — the existing IP system and its consequences are simply not suitable to solve civilizational challenges, such as COVID-19, and therefore they have to be re-evaluated. Essentially, legal talks on a TRIPS waiver provide a means of attempting to settle these matters politico-economically and may encourage the sector to volunteer information regarding its processes.[1188]

Analysis and Summary

The debates about IP and patent protection have occurred recently due to the pandemic, as a patent gives the inventors the right to stop their discoveries from being copied. However, where worldwide pandemics are concerned, given that everyone stands to lose if they are not prevented or slowed down, it is counter-productive to place obstacles in the path of vaccination. The case for facilitating access to the COVID-19 vaccines by waiving patents at this time of crisis is that while the virus is allowed to spread unchecked in various parts of the world, variants are going to emerge. This has already happened in India and many other poor nations, where the huge population and its density in some areas, together with poor infrastructure, have made containment difficult. Closing borders and preventing the movement of people is detrimental to global trade and society. Therefore, immunisation would seem to be a better response and measures should be put in place to assist the sourcing of potentially usable vaccination technology in poor nations. For this to be effective, however, vaccination needs to be implemented worldwide. Thus, for financial reasons alone, aside from the health benefits, it would seem sensible to waive the patent on pandemic vaccines.

Conclusion

In the final analysis, vaccines have a long and complex history. In the past two hundred years they have been continually developed. They have contributed to improving human health and reducing the burden of disease. Yet, there are challenges relating to the time and cost in developing new vaccines, which are necessary to help prevent the spread of new diseases and their variants. There are also challenges in relation to access to vaccines in poorer countries around the world. However, vaccines and vaccine technology enjoy a range of IP rights, including patents; while vaccines themselves cannot be patented, their formulation and delivery system can be, resulting in multiple layers of protection for pharmaceutical companies.

[1186] Thambisetty (n 134).
[1187] ibid.
[1188] ibid.

Therefore, patents can help to encourage innovation, as they enable pharmaceutical companies to generate profit, but they also represent a barrier to vaccine rollouts, especially during a pandemic. Consequently, safeguarding IP during a pandemic can cause issues with manufacture and access to vital medications, which can cause greater financial problems around the globe, as well as posing a significant threat to life, especially if variants emerge. From a cost perspective alone, aside from the moral issues, waiving patents is therefore a logical choice during a time of crisis.

Afterword

Monica Kiosseva, *Deputy Editor-in-Chief*

It is with utmost gratitude that we thank you for reaching the end of this year's Law Review. We trust that you have found pieces that have not only piqued your interest but have incited your curiosity to delve deeper into new and intriguing subject matter.

We advise you, take a moment to both acknowledge and admire that in this publication lies the cumulative work of thousands of hours of research, writing, and editing.

The quintessential space for the intellectually driven. It would not be bold to assert that this has been a fervent labour of love for academia, for authors and editors alike. We have aimed not only to be a platform for aspiring legal professionals to share their dedication, but a continuing inspiration to consider how vast, varied and fascinating the legal world is, beyond our perceived scope of interests.

If even one piece you have read has persuaded you to stop reading and think about its implications, our mission has been accomplished. Thank you for sharing in our world of legal academia.

Sincerest Regards,

The Editorial Board
Volume IV
City Law Review